Regulation of Allele Specific Gene Expression of the Polycomb Group Target Gene *PHERES1*

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich
von

Grigory F. Makarevich

aus

Russland

Promotionskomitee
Prof. Dr. Ueli Grossniklaus (Vorsitz)
Prof. Dr. Claudia Köhler
Prof. Dr. Beat Keller
Prof. Dr. Ortrun Mittelsten Scheid

Zürich, 2007

Summary

Epigenetic phenomena combining changes in the gene expression status that are not associated with direct modifications of the primary coding sequence are a hot topic of current molecular genetics. Many fundamental biological processes, including cell differentiation in multicellular organisms and, in many cases, cancer formation are regulated by epigenetic mechanisms. Despite enormous progress achieved during the last years, our current understanding of the principal molecular mechanisms causing epigenetic phenomena is far away from being complete.

Genomic imprinting, a mechanism in which the expression status of an allele depends on its parental origin, is one of most prominent examples among epigenetic phenomena. It is also associated with the unusual genetic inheritance of several human genetic diseases, including the well-studied Prader-Willi/Angelman (PWS/AS) and the Beckwith–Wiedemann (BWS) syndromes (reviewed in Soejima and Wagstaff, 2005). It makes genomic imprinting a good model to study epigenetic phenomena.

PHERES1 (PHE1) is the only gene known that is maternally imprinted and paternally expressed at early developmental stages in *Arabidopsis* (Kohler et al., 2005). Therefore, *PHE1* can be used as a model to study maternal imprinting phenomena in plants. In this work I tried to investigate the role of basic epigenetic mechanisms in genomic imprinting of *PHE1* and to discover potential imprinting control regions in the *PHE1* locus.

Based on the analysis of previous reports describing epigenetic mechanisms in animal and plant model systems, I used chromatin immunoprecipitation (ChIP) techniques to check the distribution of several chromatin modifications at the *PHE1* locus at different developmental stages. I found that trimethylated lysine at position 27 of histone H3 (H3K27me3) is the most prominent mark of chromatin at the promoter of *PHE1*. I also observed significant correlation in the distribution profiles of H3K27me3 and the FIS complex subunit MEDEA at *PHE1* locus, strongly arguing that the FIS complex regulates *PHE1* expression through covalent modifications of histones.

Based on the combined analysis of FIS complex targeting, chromatin modifications, and investigations of the imprinting status of the close *PHE1* homolog *PHERES2* (*PHE2*) we predicted that FIS PcG complex targeting is required but not sufficient for the establishment of genomic imprinting at the *PHE1* locus.

Using several independent techniques we identified that the potential imprinting control element (ICE) of *PHE1* is localized in the region downstream of *PHE1*.

Our analysis of parental *PHE1* expression in several mutants affecting different steps in establishment and maintenance of DNA methylation in *Arabidopsis*, predicts that DNA methylation is involved in control of *PHE1* expression. We have identified the sites of DNA methylation in the region located downstream of *PHE1* and we predict that these sites can be part of the ICE.

Based on the experimental data obtained in the course of this work we predict a model of imprinting control at the *PHE1* locus.

1 Introduction

1.1 Genomic imprinting and epigenetic regulation of gene expression

Genomic imprinting refers to the phenomena of mitotically and/or meiotically heritable functional non-equivalency of maternal and paternal alleles of genes or genomic regions. If an allele is under control of genomic imprinting its expression status depends on its parent of origin, i.e. the activity of the allele is determined whether it was inherited from the maternal or paternal side. The inactive allele of an imprinted gene is referred to as being imprinted.

Genomic imprinting is an *epigenetic phenomenon*, i.e. heritable changes of gene function are not associated with modifications of the primary DNA sequence. Molecular mechanisms associated with epigenetic phenomena are of special interest, because these mechanisms are supposed to regulate cell type differentiation in multi-cellular organisms and are as well associated with cancer formation (reviewed in Cavalli, 2006; Sparmann and van Lohuizen, 2006).

1.1.1 Non-random X-chromosome inactivation

Probably the best studied example of genomic imprinting is the phenomenon of non-random X-chromosome inactivation. X-chromosome inactivation is one of the mechanisms that has evolved to equalize the levels of X-linked gene transcripts between heterogametic (XY) and homogametic (XX) sexes. The choice of the X-chromosome that will be inactivated in all tissues of marsupials and in the extraembryonic tissues of some eutherians like mice is not random – the paternally derived chromosome is always chosen to be inactivated (Sharman, 1971; Takagi and Sasaki, 1975).

In contrast, the choice of the X-chromosome to be inactivated in embryonic tissues of mice is random and there is an equal probability that cells will inactivate their maternally or paternally derived X chromosome. However, detailed investigations of the phenomenon revealed that the same molecular mechanisms are involved in inactivation in both cases. Initiation of X chromosome inactivation is associated with transcription of

the long non-coding RNA *Xist* (X-inactive-specific transcript). The inactive X chromosome becomes coated with the *Xist*-RNA in *cis* along the whole chromosome. Completion of the process is associated with loss of activating and progressive acquiring of inactivating histone modifications H3K9 deacetylation, H3K4 demethylation, H3K27 trimethylation, H3K9 dimethylation, and H4K20 monomethylation (reviewed in Heard and Disteche, 2006).

Two different scenarios have been proposed to explain the non-random character of X-inactivation – either the *Xist* locus of the maternally derived chromosome is predetermined inactive, or the paternally derived chromosome contains specific activation marks. The observation that mouse embryos carrying two maternally derived X chromosomes fail to develop extraembryonic tissues due to the failure to inactivate one of the chromosomes, strongly suggests that the maternally derived X chromosome carries preset inactivation marks at the *XIST* locus (Goto and Takagi, 1998). On the other hand, mice carrying only paternally derived X chromosomes are viable and develop absolutely normal extraembryonic tissues (Papaioannou and West, 1981 citied by Heard and Disteche 2006), indicating that there is no initial commitment of the paternal X chromosome to become inactivated.

1.1.2 Locus-specific genomic imprinting in animals

While X-chromosome inactivation provides an example of imprinting at the chromosomal level, there are documented examples of genomic imprinting at the level of single genes.

Detailed molecular mechanisms regulating genomic imprinting were determined for six imprinted gene clusters in the mouse genome – *Igf2*, *Igf2r Kcnq1, Pws, Gnas* and *Dlk1*. The common feature of all imprinted gene clusters and the imprinted X-chromosome is the presence of non-coding (nc) RNAs. In all cases, silencing of genomic loci (or the whole chromosome) in *cis* is correlated with transcription of the corresponding ncRNA (reviewed in Reik and Lewis, 2005). However, the functional roles of ncRNAs in establishment and maintenance of imprinted loci seems to be different. At paternally imprinted *Igf2r* and *Kcnq1* loci transcription of *Air* and *Kcnqot1* ncRNAs, respectively, is

required although not sufficient for imprinting (Mancini-Dinardo et al., 2006; Sleutels et al., 2002). In contrast, transcription of the ncRNA *H19* is not required for silencing of the maternally imprinted *Igf2* locus (reviewed in Arney, 2003; Wolffe, 2000). Based on these observations, two different models for locus-specific imprinting in mammals were suggested – boundary-dependent and ncRNA-dependent (reviewed in Pauler and Barlow, 2006). However, even for those loci requiring transcription of ncRNAs for imprinting establishment, the precise role of ncRNAs for the imprinting mechanism is not completely understood. Furthermore, the exact mechanism of the initial differentiation between maternal and paternal alleles, causing differential transcription of ncRNAs awaits further investigations.

Regulatory models of both boundary-dependent and ncRNA-dependent imprinted loci in mammalian system are presented in Fig 1-1

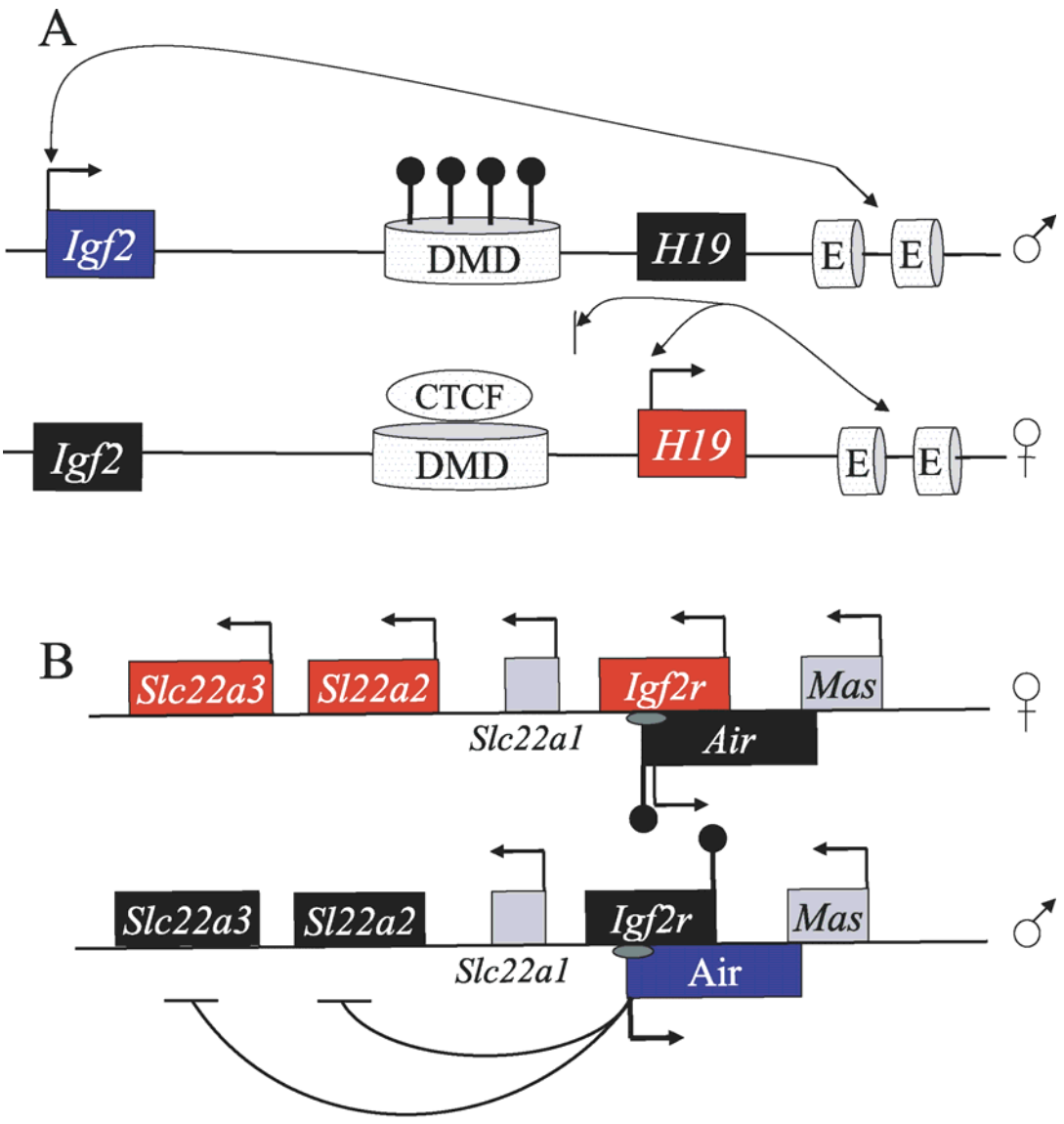


Fig 1-1 | Regulatory models at imprinted loci in mammals (with modifications from Wood and Oakey, 2006) Blue boxes represent paternally expressed alleles, red boxes maternally expressed alleles, black boxes silenced alleles, and grey boxes nonimprinted genes. Arrows on boxes indicate transcriptional orientation. (A) The boundary model is well studied at the *Igf2/H19* locus and consists of an ICR located between a pair of reciprocally expressed genes that controls access to shared enhancer elements. On the paternal allele, the differentially methylated domain (DMD) acquires methylation (black circles) during spermatogenesis, which leads to repression of the *H19* promoter. The

hypomethylated maternal DMD acts as an insulator element, mediated through binding sites for the methylation-sensitive boundary factor CTCF (shaded ellipse). When CTCF is bound, *Igf2* promoter access to the enhancers (E) distal to *H19* is blocked. (B) At the *Igf2r* locus, the paternally expressed, noncoding RNA *Air* acts to induce bidirectional cis-mediated silencing (black curved lines) on neighboring protein-coding genes (maternally expressed *Igf2r*, *Slc22a3*, and *Slc22a2*). The grey ellipses are the intronic imprint control elements that are maternally methylated (black circles) and contain the promoter of the *Air* RNA.

1.1.3 Locus-specific genomic imprinting in *Arabidopsis*

Double fertilization is a distinguishing feature of angiosperm plant development. During fertilization, the two male sperm cells fuse with the two female gametes, the egg cell and the central cell, giving rise to the embryo and endosperm, respectively. The two female gametes are genetically not identical; the egg cell is haploid and the central cell is homo-diploid. This results in genetically distinct fertilization products. Whereas in the embryo the ratio of maternal to paternal genomes is balanced (m:p = 1:1), the endosperm contains a double dosage of maternally contributed genomes (m:p = 2:1). Thus far, unambiguous evidence for genomic imprinting has only been obtained in the endosperm. The first gene shown to be regulated by genomic imprinting is the *R* gene, responsible for pigmentation of the maize endosperm (Kermicle, 1970). Recent work revealed the imprinting mechanism of the *MEDEA* (*MEA*) gene (Crossniklaus, 1998), encoding a subunit of a PcG-like complex in *Arabidopsis* (Kinoshita et al., 1999; Vielle-Calzada et al., 1999). The maternal *MEA* allele is active in the endosperm, whereas the paternal allele is inactive. It is thought, that the initial differences between maternal and paternal *MEA* alleles are generated by the DNA-glycosylase DEMETER (DME). DME acts in the central cell of the female gametophyte and removes DNA-methylation marks at regions surrounding *MEA* (Gehring et al., 2006). Activity of DME is required for expression of the maternal *MEA* allele shortly after fertilization. After fertilization, inactivation of the paternal allele is achieved through the preferential binding of MEA to the inactive,

paternal allele (Baroux et al., 2006; Gehring et al., 2006; Jullien et al., 2006b).
DNA methylation underlies the regulation of two other imprinted genes in *Arabidopsis* – *FWA* and *FIS2* (Jullien et al., 2006b; Kinoshita et al., 2004). For both genes only the maternally derived allele is active in the endosperm, the paternally derived allele is silent. Similarly to *MEA*, *FWA* and *FIS2* are methylated in all tissues. DME removes methylation in the central cell of female gametophyte causing the maternally derived alleles being unmethylated and active in the endosperm (Kinoshita et al., 2004).

1.1.4 Functional role and evolutionary origin of genomic imprinting

Analysis of the functional roles of imprinted genes in mice revealed that many of them are involved in the control of nutrient acquisition by the fetus through the placenta (reviewed in Constancia et al., 2004). Interestingly, maternally and paternally imprinted genes have opposite effects on nutrient transfer. Whereas maternally active genes reduce, paternally active genes stimulate nutrient transfer from the maternal organism to the fetus. The most remarkable example of opposing parental effects is provided by the imprinted genes *Igf2* and *Igf2r*. The maternally expressed Igf2r degrades excess of the paternally expressed growth factor *Igf2* (reviewed in Reik et al., 2003). These observations provided the basis for the widely accepted "parental conflict" hypothesis. Imprinting has only been observed in mammals and flowering plants. In both taxa maternal and paternal parents contribute unequally to the costs of the developing progeny. Haig and Westoby (1991) proposed that the interests of the maternal parent are best served if her resources are equally distributed among the siblings. In contrast, the paternal interests will favor maximal nutrient transfer, as the costs experienced by the paternal parent are negligible. These different interests create a conflict that will favor growth promoting genes being paternally active and growth suppressing genes being maternally active (reviewed in Baroux et al., 2002). Mathematical modeling supports this hypothesis, showing that selective forces can favor the origin of imprinted gene expression under such assumptions (reviewed in Wilkins and Haig, 2003). The parental conflict hypothesis is also indirectly supported by the failure to identify imprinting genes in species where there is no direct nutrient transfer from the mother to the developing

progeny, , like in fish, amphibians, reptiles, and birds. Furthermore, among mammals, only in eutherian mammals and marsupials imprinting has been reported, whereas in egg laying monotrememes imprinting has not been observed (Killian et al., 2000).

Similar to the placenta of mammals, the endosperm of angiosperms is playing an important role in the transport of resources from the maternal plant to the developing seed. Therefore, the parental conflict theory has been used to explain the evolutionary origin and functional role of imprinting in plants. Indeed, several observations strongly argue for the importance of genomic imprinting in resource distribution in plants. The most remarkable observation in favor of this theory is the seeds size dependence on maternal/paternal genome ratios. Crosses between diploid and tetraploid plants in *Arabidopsis* produce viable triploid embryos surrounded by endosperms with different maternal to paternal (m:p) genome ratios. While regular crosses of diploid plants form endosperms with a m:p ratio of 2:1; crosses of diploid mother plants with tetraploid pollen donors form endosperms with a m:p ratio of 2:2 and the reciprocal cross will produce an endosperm with a m:p ratio of 4:1. The size of seeds produced by these crosses consistently decreases with an increasing m:p ratio in the endosperm (Scott et al., 1998). One accepted explanation for these observations assumes that changes in the maternal to paternal genome ratio in the endosperm causes changes in the ratio between products of imprinted genes as well as any other dosage-sensitive gene products. This causes a shift in the equilibrium of resources transferred from the maternal plant to seeds (Scott et al., 1998). However, up to now there are only few imprinted genes known. Furthermore, the role of the known imprinted genes in nutrient transfer is unclear. Therefore, our current knowledge does not provide sufficient ground to accept or reject the parental conflict theory to explain the evolutionary origin and functional role of genomic imprinting in angiosperms.

1.2 Maternal imprinting of *PHERES1*

In contrast to many genes shown to be preferentially maternally expressed at early developmental stages (Vielle-Calzada et al., 2000), the member of the type I MADS box gene family *PHERES1 (PHE1)* is preferentially paternally expressed in seeds at the second and third day after pollination (Kohler et al., 2005). Parent of origin dependence

of *PHE1* expression was demonstrated using two independent experimental approaches. In the first approach reciprocal crosses were performed with wild-type and *phe1* mutant plants. Using specific PCR conditions no *PHE1* transcript was detectable in the *phe1* mutant. Therefore, when using the *phe1* mutant as the maternal plant, only paternally derived transcripts were detectable, whereas if *phe1* was the pollen donor only maternally derived transcript were detected. In the second approach a single nucleotide polymorphism in the *PHE1* sequence of Col and C24 accessions was used to distinguish between maternal and paternal *PHE1* alleles. Reciprocal crosses between both accessions allowed to distinguish maternally and paternally contributed *PHE1* expression in developing seeds. Both techniques demonstrated preferential expression of the paternal *PHE1* allele (Kohler et al., 2005).

PHE1 has been identified as a direct target gene of the FIS Polycomb group (PcG) complex that is strongly upregulated in *fis* mutants (Kohler et al., 2003b). By introducing the *phe1* mutation in the *mea* mutant background, parent-of-origin dependent *PHE1* expression in *mea* mutants was analyzed. The results of this experiment clearly demonstrated that only the maternal copy of *PHE1* is upregulated in *mea* mutants. As *PHE1* is directly targeted by the MEA protein (Kohler et al., 2003b), this result suggested that silencing of the maternal allele of *PHE1* depends on the specific inactivation of *PHE1* gene in the female gametophyte by the MEA-containing FIS PcG complex.

1.3 Molecular mechanisms of epigenetic phenomena - DNA and chromatin modifications

The key feature of genomic imprinting is the epigenetic difference of maternally and paternally derived alleles. Therefore, genomic imprinting depends on the mechanisms establishing such differences and maintaining them through mitotic and meiotic divisions. The molecular basis for such “hard-coded” differences is provided by DNA methylation and chromatin modifications. Special enzymatic machineries exist to establish de-novo modifications, to maintain modifications through cell divisions and to remove earlier established modifications on both DNA and chromatin levels.

1.3.1 DNA-methylation in plants

5-Methyl-cytosine is the most typical methylation modification of DNA. Intensively methylated regions of DNA are usually transcriptionally inactive. In line with this observation, the proposed functional role of DNA-methylation in plant genomes is regulation of gene activity and defense from viruses that become silenced after inserting into the genome by DNA-methylation (reviewed in Chan et al., 2005).

However, a recent study of the genome-wide DNA methylation status in *Arabidopsis* revealed that DNA methylation and transcriptional silencing do not necessarily correlate. While there is correlation between DNA methylation and silencing in transposons and repetitive elements, there is no correlation in unique genomic regions. Highly expressed and weakly expressed genes show similar low levels of DNA methylation, whereas genes having intermediate levels of transcription also have intermediate levels of DNA methylation (Zilberman et al., 2007). This analysis was complemented by a report demonstrating that genes with high methylation levels inside their transcribed parts have higher levels of constitutive expression, while genes with methylation at their promoters regions show tissue-specific expression (Zhang et al., 2006).

Based on the sequence context and underlying enzymatic machinery two types of DNA-methylation are distinguished – symmetric (occurring at CpG, CpNpG sites) and non-symmetric (occurring at CpHpH sites).

De-novo methylation activity in *Arabidopsis* is presented by DRM (domains rearranged methyltrasferase) proteins, DRM1 and DRM2 (Cao and Jacobsen, 2002). Double *drm1/drm2* mutants lack any form of *de-novo* DNA-methylation. Complete or partial loss of *de novo* DNA-methylation activity is also observed in mutants affecting small interfering RNA (siRNA) metabolism – *rna-dependent-rna-polymerase2 (rdr2); argounaute4 (ago4); dicer-like3 (dcl3)* (Chan et al., 2004). These observations suggest a possible role of RNAi-like mechanisms in guiding of DRM1/DRM2 to target sequences.

Propagation of DNA-methylation marks through cell divisions depends on the DNA methylation maintenance machinery. In a symmetric DNA methylation context, maintenance machineries can recognize hemimethylated sites formed during DNA replication. CpG methylation in *Arabidopsis* is mainly maintained by

METHYLTRASFERASE1 (MET1) activity (Kankel et al., 2003; Saze et al., 2003). Propagation of CpNpG methylation marks in *Arabidopsis* depends on the activity of the plant-specific methyltransferase CMT3 (Bartee et al., 2001; Lindroth et al., 2001). In contrast, asymmetric methylation sites probably require active targeting of DNA-methylation enzymes during each cycle of DNA replication. Observations that mutants with defects in components of the RNAi machinery have lower level of DNA methylation at CpNpG and CpHpH sites support this view (Chan et al., 2004).

The DNA methylation level is also significantly lower in several mutants that do not directly affect genes encoding DNA-methyltransferases – *DECREASE IN DNA METHYLATION1 (DDM1)* (Jeddeloh et al., 1999), and *KRYPTONITE (KYP)* (Jackson et al., 2002; Malagnac et al., 2002). DDM1 and KYP are involved in chromatin remodeling and histone methylation, respectively, indicating that chromatin modifications and DNA-methylation are functionally linked to each other.

DNA-demethylation activity in *Arabidopsis* is mediated by a small family of DNA-glycosylases (Choi et al., 2002; Gong et al., 2002). DEMETER (DME) is one member of this family that is specifically expressed in the central cell during female gametophyte development (Choi et al., 2002). Activation of maternal *MEA* and *FWA* alleles in the endosperm requires active removal of DNA-methylation marks from the maternal alleles of those genes by DME (Choi et al., 2002; Gehring et al., 2006; Kinoshita et al., 2004).

1.3.2 Chromatin modifications in plants

Association of DNA in the nucleus with proteins results in the formation of nucleoprotein complexes that provide an additional level of gene regulation through the modulation of the physical accessibility of DNA to the transcription machinery. Regulatory mechanisms include:

- Phasing of nucleosomes
- Posttranslational modifications of histones
- Substitution of histones with other proteins
- Nuclear organization at higher order

Posttranslational modifications of histones are epigenetic marks specific to active and

inactive transcriptional domains. Two types of histone modifications – acetylation and methylation – were known for many years. However, only recently evidence is accumulating pointing towards an important role of both modifications in the epigenetic regulation of gene expression. While acetylation of histones is usually associated with increased gene activity, effects of histone methylation are very specific and dependent on the specific histone and amino acid residue being modified.

Analysis of the *Arabidopsis* genome revealed at least 12 histone acetyltransferases (HATs) and 17 histone deacetylases (HDACs) (Pandey et al., 2002). Involvement of HDACs in epigenetic regulation is not limited to modification of chromatin structure; the same proteins are also directly or indirectly involved in the maintainance of DNA methylation. Thus, mutants in the *HAD6* gene demonstrate loss of DNA methylation at specific genetic loci (Aufsatz et al., 2002; Murfett et al., 2001; Probst et al., 2004). This observation is an additional argument supporting the hypothesis of functional linkage between different epigenetic mechanisms.

Several different lysine residues at histones H3 and H4 can be mono-, di-, and trimethylated. In plants, methylation of H3K9, H3K27, and H4K20 is associated with inactive and methylation of H3K4, and H3K36 correlates with active chromatin structure (reviewed in Fuchs et al., 2006).

Histone methylation (except methylation of H3K79 to date not characterized in plants) is mediated by histone methyltransferases belonging to the conserved SET-domain protein family (reviewed in Sims et al., 2003). According to plant chromatin database (www.chromdb.org), *Arabidopsis* has 41 SET-domain containing proteins. Some of them, including KRYPTONITE (KYP), MEA, CURLY LEAF(CLF), SWINGER (SWN), are characterized in details. KYP has H3K9 methyltransferase activity and is required for establishing this epigenetic mark (Jackson et al., 2002; Malagnac et al., 2002). Surprisingly, this gene was originally discovered in a mutant screen for suppressors of the epigenetic *SUP* allele, which is completely inactivated by dense DNA methylation at both symmetric and non-symmetric sites (Jackson et al., 2002). Further analysis of the functional role of KYP suggested that this protein is involved in the establishment of H3K9 methylation marks, which are recognized by the protein LHP1 (like HP1) (Zemach et al., 2006). In turn, LHP1 interacts with DNA methyltransferase

CMT3 targeting it to sites with this epigenetic mark (Jackson et al., 2002), demonstrating functional linkage of histone and DNA modifications.

MEA, CLF and SWN are homologues of the *Drosophila* PcG protein Enhancer of zeste (E(Z)). Primary targets of CLF are the homeotic type II MADS box gene *AGAMOUS* (*AG*) and *SHOOTMERISTEMLESS* (*STM*). CLF mediated repression of these loci during leaf and flower development is associated with trimethylation of lysine 27 of histone H3 (H3K27) (Schubert et al., 2006). As it was described above, MEA is expressed at early stages of seed development and is involved in regulation of *PHE1* expression (Grossniklaus et al., 1998; Kohler et al., 2003b). However, the exact molecular mechanism of MEA action is not known.

1.3.3 Role of Polycomb group proteins in epigenetic modifications

Enhancer of zeste (E(Z)) is a subunit of the Polycomb Repressive Complex 2 (PRC2), together with Suppressor Of Zeste 12 (SU(Z)12), Extra sex combs (ESC), and P55 (reviewed in Schwartz and Pirrotta, 2007). Detailed investigation of this complex in animal systems (mainly fly and mice, including cell cultures) suggested that the primary role of this complex is the establishment of H3K27 trimethylation marks over the regions to be inactivated (Czermin et al., 2002; Muller et al., 2002). Further events in establishing and maintaining the inactive state include the activity of other Polycomb Repressive Complexes (like PRC1 or PhoRC) (reviewed in Schwartz and Pirrotta, 2007; Sparmann and van Lohuizen, 2006), or recruiting of DNA-methyltransferase machinery (Vire et al., 2006)

In *Arabidopsis* several subunits of PRC2-like complexes have been identified, regulating different transitions during the plant life-cycle (reviewed in Guitton and Berger, 2005). One of these contains MEA (homolog of E(Z)), FIE (homolog of ESC), FIS2 (homolog of SU(Z)12), and MSI1 (homolog of P55) (Kohler et al., 2003a, Chanvivattana et al., 2004). MEA, FIE and FIS2 were initially discovered in a screen for fertilization independent seed (fis) development or maternal effects and mutants in these genes are characterized by similar mutant phenotypes (Grossniklaus et al., 1998; Luo et al., 1999; Ohad et al., 1999). FIS2 is the subunit that is likely to determine complex specificity, therefore, we named this complex FIS complex.

Thus far, no functional homolog of a PRC1 complex has been identified in Arabidopsis that is supposed to stabilize the inactive chromatin status initiated by PRC2. This indicates that other, as yet unknown protein complexes have this functional role in plants.

1.4 Aim of the thesis

Genomic imprinting represents one of the most interesting examples of epigenetic phenomena. While being well documented for several cases in animal and plant models, the exact molecular basis and functional significance of genomic imprinting are still poorly understood. While significant progress has been made in understanding the regulation of maternally expressed imprinted genes, only little is known about the regulation of paternally expressed imprinted genes in plants. Thus far, *PHE1* is the only reported gene being paternally expressed, making *PHE1* an interesting model to study genomic imprinting in plants. Therefore, the goal of this thesis was to investigate the molecular mechanisms regulating allele-specific gene expression of *PHE1* and to understand the role of the FIS complex in this process.

2 Materials and methods

2.1 Standard molecular biology protocols, enzymes, chemicals, and kits

2.1.1 Basic protocols and chemicals

All standard molecular biology procedures (e.g. restriction enzyme digestions, ligation reactions or plasmid DNA-preparations) were performed as described in (Sambrook et al., 1989) or according to the suppliers' instruction. Restriction endonucleases, DNA modifying enzymes and reaction buffers were purchased from New England Biolabs (USA) and used as recommended by the manufacturer. *Taq*-DNA polymerase was obtained from Sigma (USA) or was home-made following the protocol of (Desai and Pfaffle, 1995). Chemicals were purchased from Sigma-Aldrich, unless otherwise indicated.

The following kits were used for standard molecular procedures:

GFX PCR DNA and Gel Band Purification Kit (Amersham, Freiburg, Germany),
Qiagen PCR cloning kit (Qiagen, Basel, Switzerland)

2.1.2 Standard PCR reactions

Unless otherwise indicated, the following protocol was used to perform standard PCR:
Standard PCR buffer: 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$
dNTPs mix: 0.2mM each
Primers: 0.1µM each
Taq-DNA Polymerase: 1.25u/µl
DNA template: 1ng
PCRs were performed in a standard DNA thermocycler engine DYAD (MJ Research)

with the following program (94°C – 2 min; 30-45x(94°C – 15 s; 55°C – 30 s; 72°C – 45 s); 72°C – 10 min)

2.2 Plant material and growth conditions

Plants were grown in growth chambers at 70% humidity and daily cycles of 16h light at 21°C and 8 h darkness at 18°C.

The *mea* mutant used in this study was the *mea-1* allele described by (Grossniklaus et al., 1998).

The *clf* and swn mutants are T-DNA insertion mutants SALK_021003 and SALK 050195, respectively.

The *met1* mutant used in this study is *met1-3* T-DNA insertion allele described in (Saze et al., 2003). Homozygous *met1* plants were always selected among progeny of heterozygous plants and genotyped using primers MEF-1 (GATTGT-GTCTCTACTACAGAGGC) and MER-1 (GTTAAGCTCATTCATAGCCTTGC) as described in (Saze et al., 2003).

The *drm1/drm2* mutant has homozygous mutations in both *DRM1* (*AT1G28330*) and DRM2 (*AT5G14620)* genes is obtained from NASC as the line CS6366.

The *rdr2* mutant is *rdr2-2* T-DNA insertion loss-of-function mutant (SALK_059661) into the gene *AT4G11130*.

The *dcl3* mutant is *dcl3-1* T-DNA insertion mutant (SALK_005512) into the gene *AT3G43920*.

The *kyp* mutant is T-DNA insertion mutant (SALK_105816) into the gene *AT5G13960*.

The *cmt3* mutant is T-DNA insertion mutant (SALK_148381) into the gene *AT1G69770*.

SALK_023744 contains a T-DNA insertion in the 3' region of *PHE1*.

The primers used for the genomtyping assays of SALK T-DNA insertion mutants are summarized in the following table:

Mutant	Left border primer	Right border primer
rdr2 (SALK_059661)	GM13: TTCCCGCAAATGCTCTCTCTG	GM14: TTTCATCCATGCAACCGATCA
dcl3 (SALK_005512)	GM15: AGGCTGTGTGGTTCGTGTGGT	GM16: GAGAAGTAGGCATCTCTCTGCAA
kyp (SALK_105816)	GM19: GCCAGCATAAGCTTTGGTCTT	GM20: GCCCTTTTGTGCCCAGAACTT
cmt3 (SALK_148381)	GM95: TACTGGTAACGGAAGGATGCC	GM96: AGCTATGTCGACAGGGTTGTG
SALK_023744	GM276: GGTTCAATTCCAATACCAGGC	GM277: CGTATTCATGTGTCACGGTTG

If not specified, all genotyping assays were designed using T-DNA primer design tool at http://signal.salk.edu/tdnaprimers.2.html.

The *pickle* mutant used in this study is the *pkl1-1* loss-of-function allele obtained from NASC, stock number CS3840.

The *FUS3::GUS* line was provided by François Parcy (Kroj et al., 2003). To introduce the *FUS3::GUS* construct into the *mea* mutant background, the line was crossed with *mea* mutants and *mea*+/-; *FUS3::GUS/FUS3::GUS* plants were selected in the F2 generation based on the *mea* mutant phenotype and lack of selection marker segregation for the *FUS3::GUS* construct.

The transformant lines DB203 and DB215 have *GUS* gene under the control of 800bp *PHE1* promoter. These lines were generated and initially characterized by Claudia Kohler (unpublished data).

2.3 Chromatin immunoprecipitation

2.3.1 ChIP protocol

We followed with minor modifications the ChIP protocol described in (Bowler et al., 2004). If not specifically indicated all the steps are done at +4 °C. Approximately 0.5 g of material from corresponding stage was harvested into cold double distilled water and vacuum infiltrated with 1% formaldehyde for 15 min. The crosslinking reaction was stopped by vacuum infiltration with the addition of glycine to final concentration of 125 mM for 5 min. Crosslinked material was grinded to fine powder in liquid nitrogen and resuspended in 10 ml extraction buffer 1 (0.4 M sucrose,10 mM Tris-HCl pH 8, 10 mM MgCl2, 0.1 M PMSF and Complete® Protease inhibitor (Roche)). The samples were then filtered twice through Miracloth® (Calbiochem) and centrifuged for 20 min at 3000g. The resulting pellet was resuspended in extraction buffer 2 (0.25 M sucrose,10 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 1 % Triton X-100, 0.1 M PMSF and Complete® Protease inhibitor) followed by a centrifugation step for 10 min at 12000g. The pellet was resuspended in 400 µl extraction buffer 3 (1.7 M sucrose,10 mM Tris-HCl pH 8, 2 mM $MgCl_2$, 0.15 % Triton X-100® and Complete® Protease inhibitor) layered on the top of 400 µl of the same extraction buffer 3, then centrifuged for 1 h at 16000g. Pellet containing purified nuclei was resuspended in 300 µl of lysis buffer (50 mM Tris-HCl pH8, 10 mM EDTA, 1 % SDS and Complete® Protease inhibitor). The chromatin was then sonicated under conditions, practically adjusted to generate the DNA fragments ranging in size from 0.3 kb to 1.5 kb (once being established the same settings were used for all experiments). After sonication, chromatin was diluted 10 times by adding ChIP dilution buffer (16.7 mM Tris-HCl pH 8, 1.2 mM EDTA, 1.1 % Triton X-100®, 167 mM NaCl and Complete® Protease inhibitor); the solution was then centrifuged for 5 min at 4500g and the supernatant was split into 1 ml aliquots. Chromatin samples were pre-cleaned with 40 µl of salmon sperm-sheared DNA/protein A agarose beads for 1 hour at 4 °C with gentle agitation; centrifuged at 16000g for 2 min to pellet beads. Supernatant containing pre-cleaned chromatin was incubated with gentle agitation overnight with corresponding antibodies. Aliquots of 25 µl were collected at this stage to represent input

fraction. Immunoprecipitates were collected by incubation with gentle agitation with sperm-sheared DNA/protein A agarose beads for 1 hour; then beads were precipitated and washed with the following buffers: (i) Low salt wash buffer (150 mM NaCl, 0.1 % SDS, 1 % Triton X-100, 2 mM EDTA, 20 mM Tris-HCl) twice. (ii) High salt wash buffer (500 mM NaCl, 0.1 % SDS, 1 % Triton X-100, 2 mM EDTA, 20 mM Tris-HCl) twice. (iii) LiCl wash buffer (250 mM LiCl, 1 % Nonidet 40, 1 % desoxycholate-Na, 1 mM EDTA, 10 mM Tris-HCl) twice. (iv) TE buffer twice.

Chromatin was eluted from beads by incubation with elution buffer (1 % SDS, 0.1 M $NaHCO_3$) at 65 °C followed by beads precipitation by centrifugation at 4500g for 2 min. Supernatant containing chromatin was reverse cross-linked by adding NaCl to final concentration of 0.2M and incubating overnight at 65 °C.

After proteinase K treatment over 2h at 45 °C, DNA was extracted by phenol/chloform procedure and precipitated by isopropanol with sodium acetate in the presence of 20 µg of glycogene as DNA carrier. Precipitated DNA was washed with 75% ethanol, dried and distilled in double distilled water.

2.3.2 Antibodies

anti-MEA antibodies were described in (Kohler et al., 2003b) supplementary material. Antibodies against H3K27me2 were kindly provided by Thomas Jenuwein, antibodies against other histone modifications (H3K27me3, H3K4me2, H3K9me2) were purchased from Upstate (Charlottesville, USA).

2.3.3 Quantitative real time PCR analysis of precipitated material

3µl of ChIP-precipitated and input DNA samples were used to amplify regions of *PHE1*(*AT1G65330*) and *UBQ11*(*AT4G05050*) using 1× SYBR green PCR master mix (Applied Biosystems) and 12.5 pmol of each gene specific primer in a total volume of 25 µl. The amplification conditions were 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C, and 60 s at 60°C. The specificity of the unique amplification product was

determined by melting curve analysis according to the manufacturer's instructions. Each reaction was done in triplicate. For ChIP-precipitated and input fractions the ratio of averaged over triplicates absolute levels of *PHE1* to *UBQ11* was calculated. Relative enrichments were calculated as the ratio of the obtained values in immunoprecipitated and input fractions, according to Geisberg JV, Struhl K (2004) Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. In Aparicio O (ed.), Curr Prot Cell Biol pp 17.7.1-17.7.23, John Wiley and Sons, Inc, USA. All ChIP experiments were repeated at least once.

The primers used for quantitative real time PCR analysis of ChIP experiments are summarized in the following table:

Amplification Region	Primers
Region 1 *PHE1*	GM154: TGATGTTTAGTTACTAATATTTTTTTTGGCA
	GM155: TGCTATTATGTGTGTTATCGATTTTGG
Region 2 *PHE1*	GM156: AGAAAATGATTCAGTGAGGAAAACAA
	GM157: CGGATGACCGCACATGC
Region 3 *PHE1*	GM158: TTGGTGTAGCTCCTACTGTTGTGG
	GM159: GAATTGAACCGGCTCTTGCT
Region 4 *PHE1*	GM160: CATCAATGATCTCTAGTCGAAGGGA
	GM161: CTGTAGTACAATAATAAAAAAAAACATCACATAAA
p*PHE1*::*GUS*	GM152: ATAAACCATCCTCCTCATGCTAA
	GM153: TCAGTTTAAAGAAAGATCAAAGCTC
UBQ11 (AT4G05050)	GM150: GCAGATTTTCGTTAAAACC
	GM151: CCAAAGTTCTGCCGTCC

2.3.4 Semi-quantitative ChIP PCR analysis

The number of cycles for semi-quantitative PCR from ChIP-precipitated material was empirically determined to achieve linear reliable PCR product signals. PCR products were run on 2% agarose gel in 0.5x TBE buffer. Images of the gels were captured using Intas gel imaging system (Intas).

The primers used for the analysis of enrichment of promoter regions of *PHE1* and *PHE2* are summarized in the following table:

Amplification region	Primers	Amplicon size (bp)
PHE1 (At1G65330)/ PHE2 (At1G65300) promoters	CK108: CACATGCCTACACGTAAG	*PHE1*: 410
	CK167: TCCAACACCGAAAACTCCAT	*PHE2*:383
ACTIN7 (*AT5G09810*)	OL60: CGTTTCGCTTTCCTTAGTGTTAGCT	181
	OL62: AAATCGGCATAGAGAATCAA	

2.4 RNA extraction and quantification of messenger RNA levels

2.4.1 RNA isolation

RNA isolation was performed using TRIzol® (Invitrogen) following manufactures protocol. Three siliques or emasculated flowers were used for RNA isolation for time courses analysis. Three to five mature leaves were used for RNA isolation from leaves. Thirty to fifty young seedlings were used for RNA isolation for Northern-blot analysis.

2.4.2 Reverse transcription

Half of the total RNA was treated with 5 units of RNase-free DNase (Amersham Pharmacia Life Science) for 30 min. Samples were extracted with phenol-chloroform and precipitated with ethanol. The RNA was reverse transcribed using oligo-dT primer (Invitrogen) in case of gene expression analysis or using region-specific primer in case of search for noncoding RNA using Superscript II reverse transcriptase (Invitrogen) in accordance with manufactures recommendations.

Complete digestion of DNA during DNase treatment step as well as efficiency of reverse transcription was controlled by PCR from the yield of reverse transcription reaction using

primers amplifying products of different size from cDNA and gDNA.

Besides oligo-dT, the following primers were used for the first strand synthesis step of reverse transcription in the experiments analyzing existence of noncoding RNAs in *PHE1* locus:

Experiment:	Primer:
Search for antisense RNA in *PHE1* coding region – experiment	CK166: CATCCGTAGCCCGTACAACT
Search for antisense RNA in *PHE1* coding region – positive control	CK91: GAAGAAGACTCACCGTTCTCC
Search for noncoding RNA in *PHE1* downstream region – experiment	GM288: ATCAGAGTATGAATTTGGAT
Search for noncoding RNA in *PHE1* downstream region – negative control	GM290: AAATACATCAACGGAAAT

The products of first strand synthesis in these experiments were analyzed by PCR using primers combinations CK166/CK175 (CGTCTCTTGATCCACCATCTTCTTGGTCC) for *PHE1* coding region and CK288/CK290 for downstream region.

Primers used to perform PCR to control reverse transcription experiments are summarized in the following table:

Tissue	Gene	Primers:	Amplicon size (bp)	
			cDNA	gDNA
Siliques	*ACTIN11* (*AT3G12110*)	GM49: AACTTTCAACACTCCTGCCATG GM50: CTGCAAGGTCCAAACGCAGA	169	324
Leaves	*ACTIN1* (*AT2G37620*)	OL48: GGCGATGAAGCTCAATCCAAACG OL49: GGTCACGACCAGCAAGATCAAGACG	390	490

2.4.3 Quantification of messenger RNA levels

1/20th of the cDNA samples were used to amplify cDNA of gene of interest (*PHE1*(*AT1G65330*), *PHE2*(*AT1G65300*) or *MEA*(*AT1G02580*)) and control cDNA (*ACT11*(*AT3G12110*) for siliques or *GAPDH*(*AT3G26650*) for leaves) using 1× SYBR green PCR master mix (Applied Biosystems) and 12.5 pmol of each gene specific primer in a total volume of 25 µl. PCR was performed on a ABI Prism 7700 Sequence Detection System (Applied Biosystems). The amplification conditions were 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C, and 60 s at 60°C. The specificity of the unique amplification product was determined by melting curve analysis according to the manufacturer's instructions. Three replicates were performed for each sample. The ratio of averaged over triplicates absolute levels of gene of interest to control gene was calculated and presented as relative level of mRNA of gene of interest.

Primers used for the RT-PCR analysis are summarized in the following table:

PHE1 (AT1G65330)	Fwd: CGCATGTGCGGTCATCC
	Rev: TCCAACACCGAAAACTCCAT
PHE2(*AT1G65300*)	Fwd: GCCAAAGAAAAAGAGCAGCTG
	Rev: TGCATTCGCAACAATAGGAAG
MEA (AT1G02580)	Fwd: GGTGAGGCACTAGAATTGAGCAGT
	Rev: CCATAGTCCTGCCCAACCG
ACTIN11 (AT3G12110)	Fwd: GGAACAGTGTGACTCACACCATC
	Rev: AAGCTGTTCTTTCCCTCTACGC
GAPDH (AT3G26650)	Fwd: CTCCCTTGGAAGGAGCTAGG
	Rev: TTCTTGGCACCAGCTTCAAT

2.5 *PHE1* and *PHE2* imprinting assay

2.5.1 *PHE1* imprinting assay

The procedure described in (Kohler et al., 2005) was used to perform genomic imprinting assays for *PHE1*. Specifically, 1/20 of the cDNA samples were used to amplify *PHE1* using primers PHE1s2 (CGCATGTGCGGTCATCC) and PHE1as2 (CGTCTCTTGATCCACCATCTTCTTGGTCC). The amplification conditions for *PHE1* were 2 min at 94°C, 35 cycles of 15 s at 95°C, 15 s at 54°C and 30 s at 72°C, followed by 10 min at 72°C. PCR reaction products were purified using GFX kit and digested with *Hph*I restriction endonuclease for 1 hour. The digest products were separated on 2.5% agarose gel.

2.5.2 *PHE2* imprinting assay

Similar procedure was used to perform genomic imprinting assays for *PHE2*. Specifically, 1/20 of the cDNA samples were used to amplify *PHE2* using primers GM63 (CTGCATTCGCAACAATAGGAAG) and GM83 (GGAAGGCGTTGAAGACGT). The amplification conditions for *PHE2* were 2 min at 94°C, 35 cycles of 15 s at 95°C, 15 s at 55°C and 30 s at 72°C, followed by 10 min at 72°C. PCR reaction products were purified using GFX kit and digested with *Taq*I restriction endonuclease for 1 hour. The digest products were separated on 2.5% agarose gel.

2.6 Nothern- and southern-blot analysis

2.6.1 Materials

Non-radioactive Roche (© Roche Diagnostics Corporation) DIG system was used to perform both southern and northern-blot analysis in accordance with manufacturer recommendations and protocols. All required materials were ordered from Roche Diagnostics Corporation.

2.6.2 Southern-blot protocol

DNA from closed flowers was isolated using Phytopure kit (Amersham) in accordance with manufacturer recommendations. 5 µg of genomic DNA were digested overnight with 20 units of *Cla*I restriction endonuclease. Digestion products were separated on 0.8% agarose gel and transferred on Roche positively charged nylon membrane in accordance with capillary transfer method protocol from Roche. DNA was crosslinked to membrane using UV-method protocol from Roche.

2.6.3 Nothern-blot protocol

RNA was isolated from seedlings following the procedure described in section 2.4.1 of this thesis. Approx. 20 µg of RNA samples were separated on 1% agarose gel and transferred on Roche positively charged nylon membrane in accordance with capillary transfer method protocol from Roche. RNA was crosslinked to membrane using UV-method protocol from Roche.

2.6.4 Probe generation

Probes used for hybridization in both southern and northern blot analysis were generated and labeled in accordance with DIG DNA labeling by PCR protocol from Roche Diagnostics.

The mixture of two probes (*PHE1*.downstream.1 and *PHE1*.downstream.2) was used for southern-blot analysis of DNA methylation in the region located downstream of *PHE1*. The primers GM180 (TTCAATTCCAATACCAGGCT) and GM181 (GGATAAAATAGATAATACAAGCAC) were used to generate the probe *PHE1*.downstream.1 and the primers GM182 (TCTCCAAAGAGTAAACCGTA) and GM183 (TCAGTTGTAAATGACACCAG) were used to generate the probe *PHE1*.downstream.2. Both probes were generated using genomic DNA of *Col* wt as

template.

Three probes were used to perform northern-blot analysis of transcription activity on the left and right borders of T-DNA insertion in SALK_023774 - SALK_023774.left, SALK_023774.right, and SALK_023774.kan (from kanamycin ORF as a positive control). The primers GM298 (GTTGATGTATTTGCCTGA) and GM299 (ACATGACATAGCCATAGC) were used to generate the probe SALK_023774.left; the primers GM300 (CCCCACGAAGTTAGTAGT) and GM301 (GGATAAAATAGATAATACAAGC) were used to generate the probe SALK_023774.right; the primers GM302 (CTCTGATGCCGCCGTGTT) and GM303 (CGATGTTTCGCTTGGTGGTC) were used to generate the probe SALK_023774.kan. All three probes were generated using genomic DNA of the line SALK_023774 as the template.

2.6.5 Hybridization and detection

Hybridization and detection were performed in accordance with the Roche Diagnostics protocol for southern- and northern-blots hybridization and detection. Hybridization was performed at 42°C.

2.7 Bisulphite sequencing.

2.7.1 DNA isolation

DNA samples from tissues were isolated using Phytopure kit (Amersham) in accordance with manufacturer recommendations.

2.7.2 Protocol

Analysis of DNA methylation in the *PHE1* downstream region was performed in

accordance with the protocol described in (Clark et al., 1994).

2 µg of DNA were digested with *EcoR*I restriction endonuclease overnight. DNA was extracted by phenol/chloroform method, precipitated with 1/10 volume of 5 M ammonium acetate and 2 volumes of ethanol at -85°C for 15 min, washed from salts with 70% ethanol, dried out and dissolved in 50 µl of TE buffer. Bisulphite conversion reaction was performed as described in (Clark et al., 1994). Removal of free bisulphite after the end of conversion reaction was performed by purification of the samples with desalting column (Promega Magic DNA Clean-Up System).

2.7.3 Amplification, cloning, sequencing

PCR reaction was used to amplify target sequence with the primers GM250 (TATTTATATTAAAGAATTAAAAATAGTAAA) and GM251 (TCAATTATAAATAACACCAATTCAATATAA) using bisuphite treated DNA as a template. The amplification conditions for *PHE2* were 2 min at 94°C, 45 cycles of 15 s at 94°C, 30 s at 50°C and 1 min at 72°C, followed by 10 min at 72°C. PCR reaction products were separated on 1% agarose gel, purified using GFX kit, and cloned with QIAGEN PCR Cloning kit.

Plasmid DNA from individual clones was isolated using standard alkaline lysis method and additionally purified with GFX kit.

Sequencing of the individual clones was performed by Microsynth AG using the primer GM250.

2.8 GUS expresson analysis

For histochemical observation of *GUS* activities in the transgenic plants, siliques were vacuum infiltrated for 15 min. with staining buffer [50 mM Na_2HPO_4/NaH_2PO_4 buffer, pH 7.0/1 mM $K_3Fe(CN)_6$/1 mM $K_4Fe(CN)_6$/0.1% Triton X-100/1 mg/ml X-Gluc], incubated overnight at 37°C in the same buffer, and cleared with a mixture of

chloralhydrate, glycerol, and water [per 100ml: 66.7g Chloralhydrate, 8.3ml glycerol, 25ml H_2O]. Seeds were isolated from carpels and mounted on slide glasses, observed under a microscope (Zeiss Axioplan 2 imaging), and photographed with a camera (Zeiss AxioCam HRC).

3 Results

3.1 Analysis of chromatin modifications at the *PHE1* locus

Detailed analysis of the molecular basis of genomic imprinting in previously studied imprinted loci in animals demonstrated the importance of chromatin modifications in the establishment and maintenance of differential transcription states of maternally and paternally derived alleles (reviewed in Lewis and Reik, 2006). Among posttranslational histones modifications, methylation of lysine 9 and lysine 27 at histone H3 was shown to be associated with transcriptionally inactive alleles (Lewis et al., 2004; Mager et al., 2003; Umlauf et al., 2004). Multiple observations from different models strongly argue for the leading role of histone methyltransferase activity of the PRC2 PcG complex in the establishment and maintenance of these chromatin modifications. There is a close association and dependence between methylation of H3K9 and H3K27 and binding of components of the PRC2 PcG-complex, Eed and Ezh2, over the paternally derived *Kcnq1* locus in extraembryonic tissues of mice (Lewis et al., 2004; Umlauf et al., 2004). Similarly, there is a correlation between the location of these specific histone methylation marks and binding of the PRC2 PcG complex with the imprinted X chromosome (reviewed in Heard, 2004). On the other hand, X-chromosome inactivation as well as the imprinting control of some of the genes from the *Kcnq1* locus is disrupted in the absence of Eed (Mager et al., 2003; Wang et al., 2001).

PHE1 has been shown to be a direct target gene of the FIS complex and to be regulated by genomic imprinting (Kohler et al., 2003b; Kohler et al., 2005). Therefore, it was hypothesized, that the mechanisms regulating *PHE1* expression may involve chromatin modifications over the locus, and specifically posttranslational methylation of histones through histone-methyltransferase activity of plant PRC2-like PcG complexes.

3.1.1 The *PHE1* locus contains H3K27 trimethylation marks

Analysis of the distribution of histone modifications associated with the inactive state of chromatin in *Arabidopsis* revealed that H3K27me3 primarily localizes to euchromatin, H3K27me2 localizes to both euchromatin and heterochromatin, whereas H3K9me2

localizes almost exclusively to heterochromatin (Lindroth et al., 2004). Based on the euchromatic localization of *PHE1*, we predicted that H3K27me3 is the mostly prominent mark associated with the inactive state of the maternally derived *PHE1* allele. Harvesting sufficient material for chromatin immunoprecipitation (ChIP) experiments from male/female gametophytes and endosperm/embryo specific tissues is a highly laborious process. Therefore, we asked the question whether we could detect H3K27me3 methylation at the *PHE1* locus without distinguishing between maternal and paternal *PHE1* alleles. We analyzed four different regions of the *PHE1* locus (Fig. 3-1) using flowers harvested before pollination, open pollinated flowers, siliques at 2-3 days after pollination (DAP), and siliques at 4-6 DAP. In agreement with our prediction, we found strong enrichment for H3K27me3 marks in region 2 overlapping with the start-codon of the *PHE1* gene, accompanied by a weaker, but significant enrichment for this mark over nearby regions. An enrichment profile during reproductive development showed the highest enrichment peak in closed not pollinated flowers, followed by a weaker peak in siliques at 2-3 DAP (Fig. 3-2).

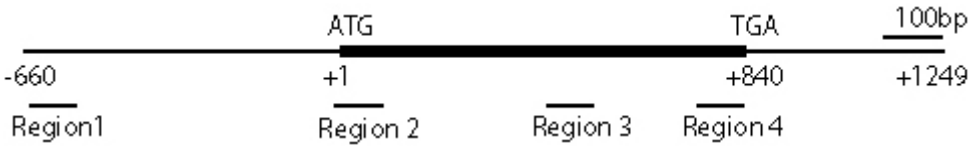


Fig 3-1 | Schematic diagram of the *PHE1* locus indicating the regions analyzed by RT-PCR after ChIP. The numbers indicate the position relative to the translational start site of *PHE1*.

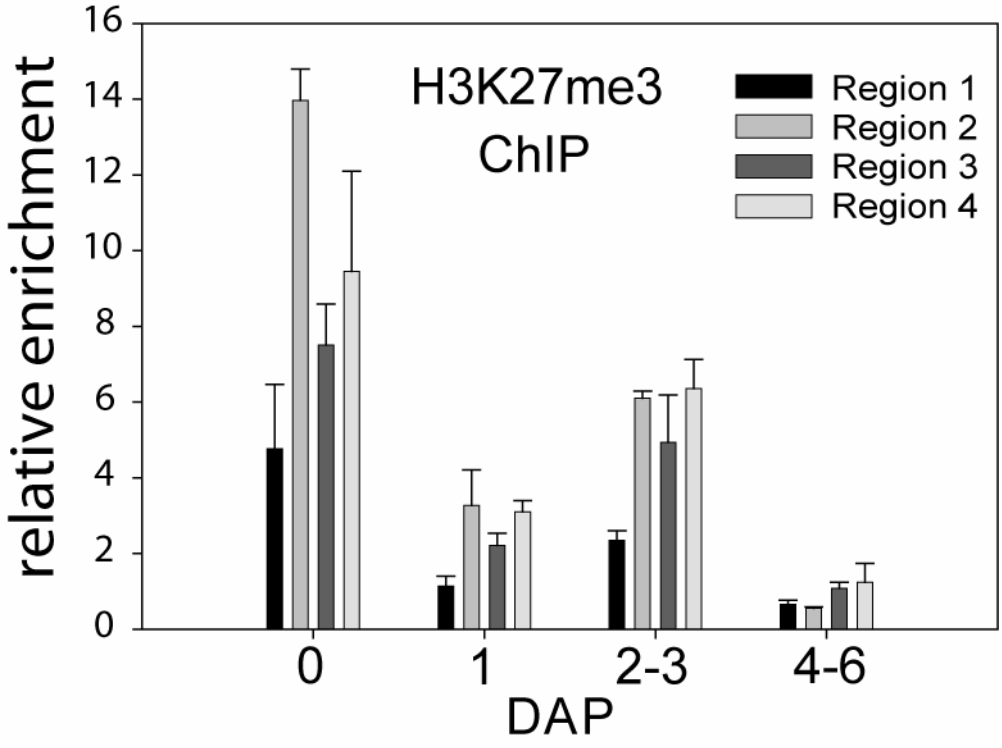


Fig 3-2 | The *PHE1* locus is enriched for H3K27me3 marks. ChIP samples with H3K27me3 antibodies were analyzed by realtime PCR in triplicate. Material used: closed flowers for 0 DAP, open flowers for 1 DAP, and siliques for 2-6 DAP.

We tested whether other chromatin modifications known to play a role in transcriptional regulation are also present at the *PHE1* locus. ChIP assays were performed using antibodies recognizing H3K27me2, H3K9me2, and H3K4me2 and material harvested from closed not pollinated flowers. We analyzed enrichment of these epigenetic marks at region 2 of *PHE1* locus and detected strongest enrichment for the H3K27me3 mark. We did not detect any enrichment for H3K27me2 inactivation mark as well as for H3K4me2 activation mark. However, we observed weak enrichment for H3K9me2 inactivation mark (Fig. 3-3).

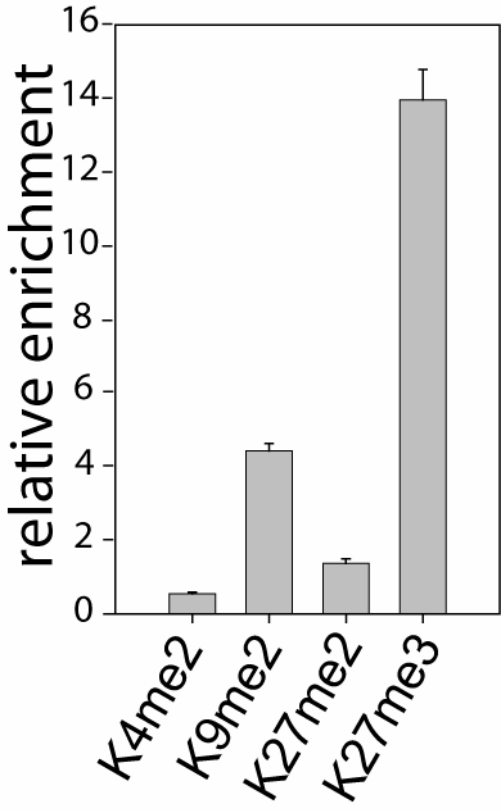


Fig 3-3 | H3K27me3 is the dominating mark present at region 2 of *PHE1*. ChIP was performed with closed flowers using antibodies targeting different modifications of histone H3. ChIP samples were analyzed by realtime PCR in triplicate.

3.1.2 Dynamics of chromatin modifications at the *PHE1* locus during development are correlated with the expression profile of *MEA* as well as with the binding of MEA to promoter regions of *PHE1*

We addressed the question whether the PRC2-like FIS-complex in *Arabidopsis* can regulate expression of its target gene *PHE1* through methylation of histone H3 at K27 by the predicted histone methyltransferase activity of MEA. Therefore, we analyzed the *MEA* mRNA profile as well as the binding profile of MEA to the *PHE1* locus and compared these profiles with the enrichment profile for H3K27me3 marks. We discovered a remarkable correlation of all three profiles and detected the presence of two peaks: one before fertilization and the other 2 days after fertilization (Fig. 3-4).

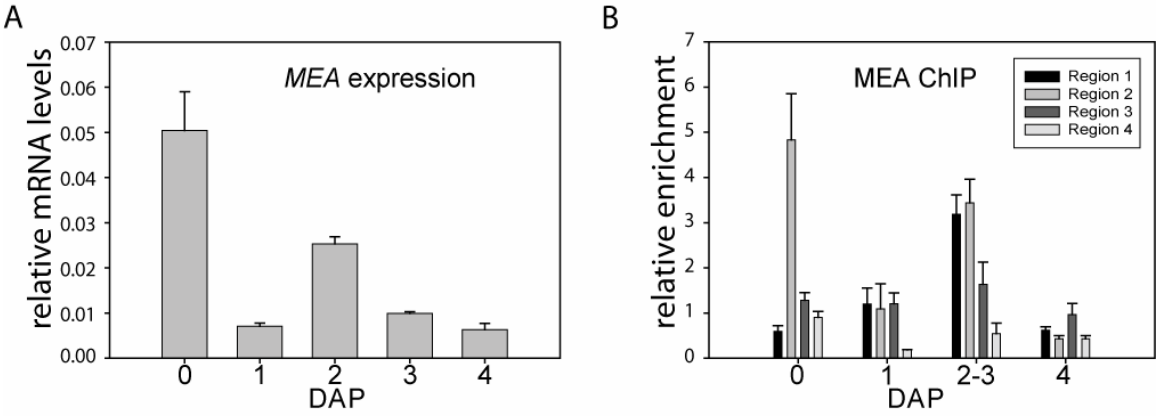


Fig 3-4 | Dynamics of H3K27me3 chromatin modification at the *PHE1* locus during development (Fig 3-2) is correlated with the expression profile of *MEA* as well as with the binding of MEA to promoter regions of *PHE1*. (A) Relative *MEA* mRNA levels in wt flowers before fertilization (0 DAP) and siliques harvested at different days after pollination (DAP). (B) Analysis of MEA binding to the *PHE1* locus during reproductive development. ChIP samples with anti-MEA antibodies were analyzed by realtime quantitative PCR in triplicate.

The close association of region- and developmental stage-specificity of MEA binding and H3K27me3 enrichment, combined with the previously reported role of the PRC2 complex in gene silencing through H3K27me3 (Cao and Zhang, 2004), strongly argues for the hypothesis that the FIS complex regulates *PHE1* expression through H3K27me3. To further test this prediction, we have compared H3K27me3 enrichment at region 2 of the *PHE1* locus in closed flowers of wild type and *mea/mea* mutants. In agreement with our prediction, we observed a decrease in H3K27me3 enrichment in homozygous *mea* mutants. This result strongly supports the prediction that MEA mediates trimethylation of H3K27 at *PHE1* (Fig. 3-5).

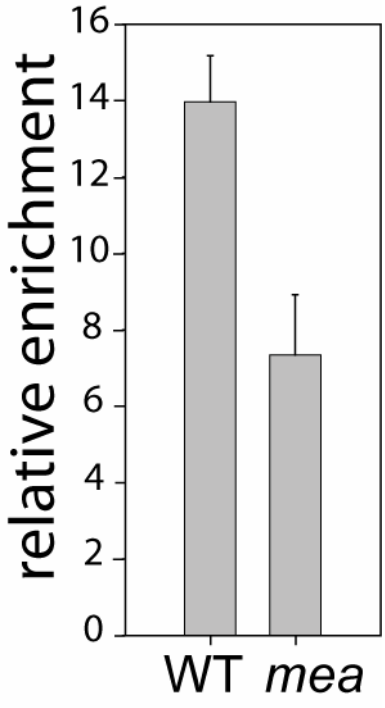


Fig 3-5 | H3K27me3 marks at region 2 of *PHE1* are reduced in *mea/mea* mutants. ChIP was performed with closed flowers using anti-MEA antibodies and samples were analyzed by realtime PCR in triplicate

3.1.3 CLF and SWN redundantly regulate *PHE1* expression in sporophytic tissues

Despite the significant decrease compared to the wild type, the enrichment level of H3K27me3 in closed flowers of homozygous *mea* mutants is still high, which argues that MEA is not exclusively responsible for histone methylation at the *PHE1* locus. As *MEA* is not expressed in sporophytic tissues (leaves) (Kiyosue et al., 1999), we predicted that the residual level of the H3K27me3 mark in closed flowers of homozygous *mea* mutant plants corresponds to the chromatin from sporophytic tissues surrounding the female gametophyte. Therefore, H3K27me3 at the *PHE1* locus is likely to be established in a MEA-independent way. To test our hypothesis, we performed ChIP experiments analyzing the abundance of H3K27me3 at the *PHE1* promoter region in chromatin extracted from leaves. In complete agreement with our prediction, we observed

significant enrichment of H3K27me3 at the *PHE1* promoter in leaves (Fig. 3-6).

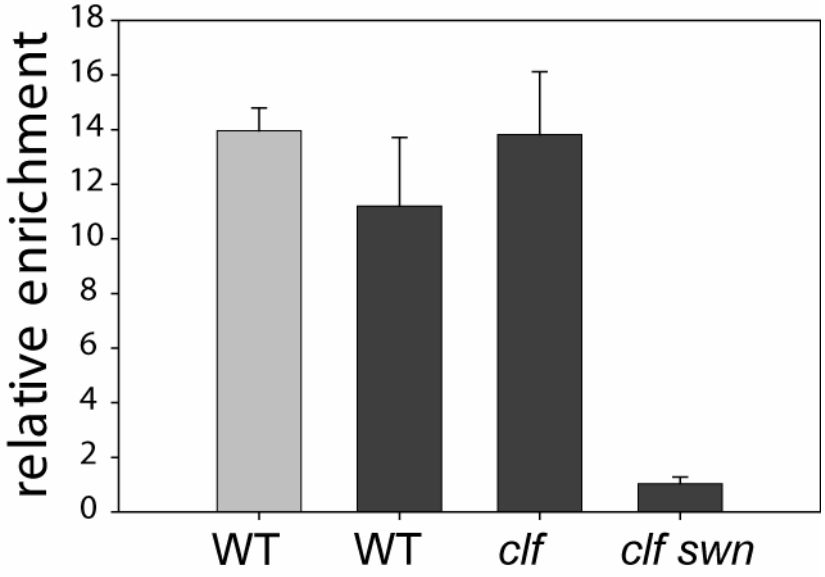


Fig 3-6 | H3K27 trimethylation of *PHE1* in sporophytic tissues is mediated by CLF and SWN. ChIP analysis of wild-type (wt), *clf* and *clf/swn* mutants was performed with anti-H3K27me3 antibodies and enrichment for region 2 of *PHE1* was analyzed by realtime PCR in triplicate. Light grey bars and dark grey bars indicate samples of closed flowers and leaf samples, respectively

CLF and SWN are homologues of MEA acting during sporophytic development. Therefore, we tested whether H3K27 me3 in the *PHE1* promoter region in sporophytic tissues depends on CLF and/or SWN activity. We performed ChIP assays from leaves of *clf* mutants, *mea/MEA*; *swn/swn* mutants and calli of *clf/swn* double mutants and analyzed H3K27me3 levels at *PHE1* region 2. While no significant change in H3K27me3 enrichment levels were observed in leaves of *clf* single (Fig. 3-6) and *mea/MEA; swn/swn* double mutants compared to wild type (data not shown), *clf/swn* double mutants demonstrated significantly lower levels of H3K27me3 at the *PHE1* locus . Thus, CLF and SWN are redundantly required for trimethylation of H3K27 at *PHE1* in sporophytic tissues. To analyze whether H3K27me3 is required for silencing of *PHE1* in sporophytic tissues, we compared the relative expression level of *PHE1* in leaves of wild-type plants;

clf and *swn* single mutants, as well as in calli of *clf/swn* double mutants. While no expression of *PHE1* was detected in wild-type, *clf* and *swn* single mutants, we observed upregulation of *PHE1* expression in calli of *clf/swn* double mutants (Fig. 3-7). However, relative expression level in *clf/swn* double mutant was still significantly lower compared to wild-type siliques, suggesting that in addition to the loss of H3K27me3, additional activating signal are required to completely activate *PHE1* expression.

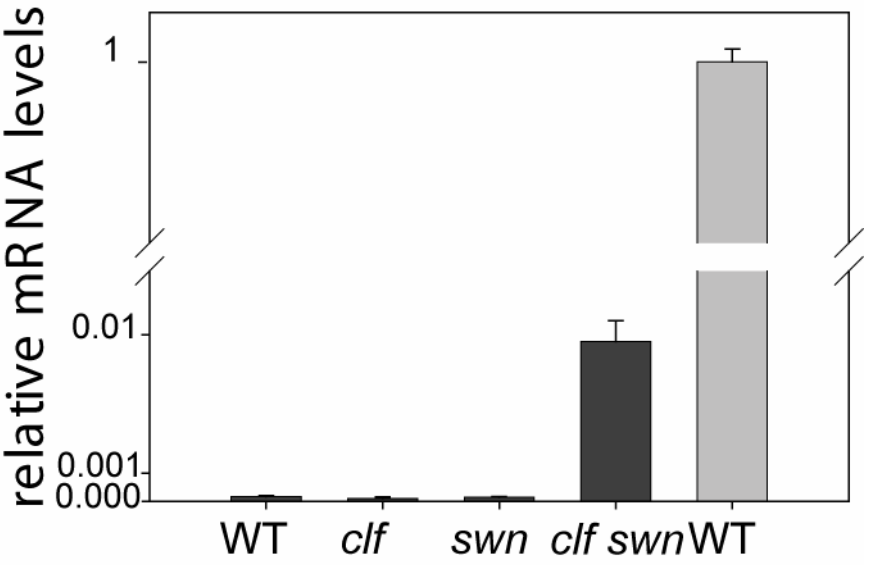


Fig 3-7 | CLF and SWN repress *PHE1* expression in sporophytic tissues. Relative *PHE1* mRNA levels were analyzed in wt, *clf* and *clf/swn* mutants. *PHE1* mRNA levels in wt seeds at 3 days after pollination indicate maximum *PHE1* expression. Light grey bars and dark grey bars indicate samples of seeds and leaf samples, respectively

3.1.4 Different PRC2-like PcG complexes in *Arabidopsis* regulate common target genes

The presence of several PRC2-like PcG complexes in *Arabidopsis* raises the question whether functional specialization of different complexes affects only tissue specificity or target specificity as well. Our experiments demonstrated that *PHE1* is targeted by FIS as well as CLF and SWN-containing complexes. Therefore, we asked whether other genes are repressed by different PcG-complexes as well.

To address this question we analyzed expression of *AG* and *AP3*, which were previously

shown to be repressed by CLF and SWN (Goodrich et al., 1997; Schubert et al., 2006). As expected, we detected a significant upregulation of both *AG* and *AP3* in *clf/swn* mutants. But we failed to detect upregulation of *AG* and *AP3* in *mea* mutant seeds.. However, as both genes are already highly expressed in seeds at 3DAP, loss of MEA function might not impact on *AG* and *AP3* expression (data not shown).

Double *clf/swn* mutants form somatic embryos (Chanvivattana et al., 2004; Lotan et al., 1998), therefore, we predicted that embryonic regulators such as *LEAFY COTYLEDON1* (*LEC1*), *LEC2* and *FUSCA3* (*FUS3*) are misexpressed in this mutant background. To test this hypothesis, we measured relative expression levels of these genes in *clf/swn* mutants. Indeed, *FUS3* was highly upregulated, while *LEC2* and *LEC1* were mildly upregulated in this background (Fig. 3-8). Thus, the embryonic regulators *LEC1*, *LEC2*, and *FUS3* are potential target genes of CLF and SWN.

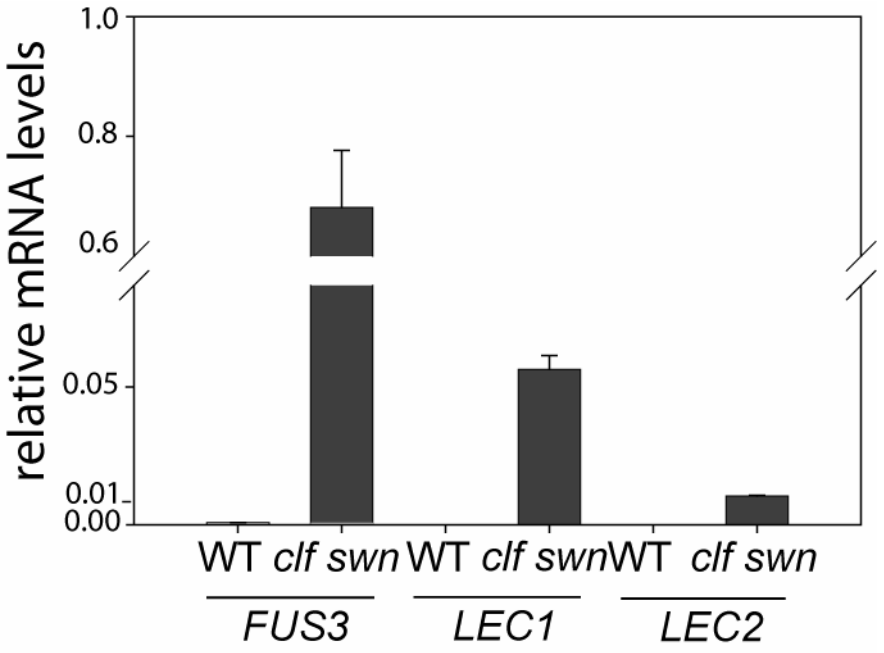


Fig 3-8 | *FUS3*, *LEC1* and *LEC2* are upregulated in *clf swn* mutants. Relative mRNA levels of *FUS3*, *LEC1* and *LEC2* in *clf/swn* mutants in comparison to wild-type (wt) leaves.

We analyzed whether these genes are also targeted by MEA by measuring their

expression levels in siliques of *mea/MEA* plants at 2-3DAP. We observed upregulation of *FUS3* expression in *mea/MEA* background (Fig. 3-9), but failed to detect upregulation of *LEC1* and *LEC2* (data not shown).

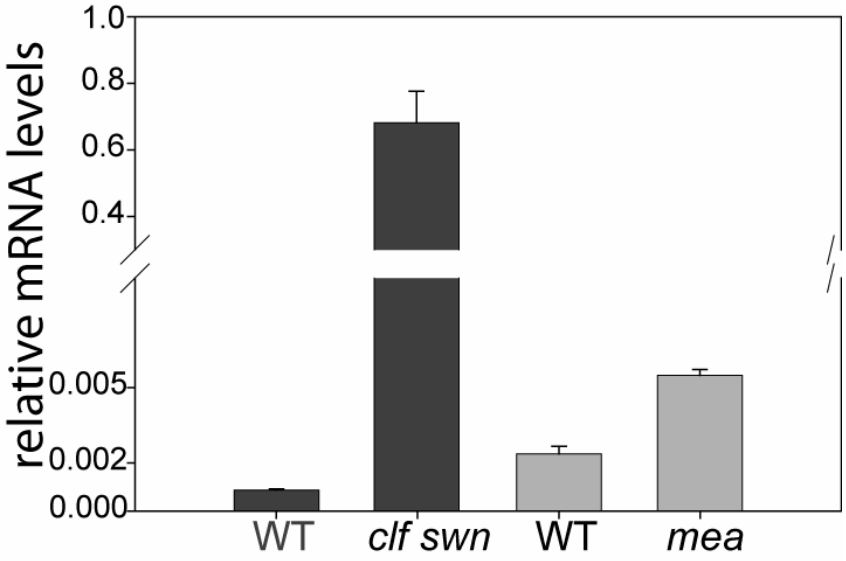


Fig 3-9 | *FUS3* is upregulated in *mea* mutants. Relative mRNA levels of *FUS3* in seeds of wild-type and *mea/+* mutants at 2-3 days after pollination (DAP). Expression of *FUS3* in *clf/swn* and wt leaves is shown as a comparison.

We also compared the expression pattern of *FUS3* and *LEC2* genes in *mea/MEA* and wild-type background using *FUS3::GUS* and *LEC2::GUS* reporter constructs. While in wild-type seeds expression of *FUS3::GUS* was restricted to the embryo until the heart stage, in *mea* mutant seeds GUS expression was detectable in the endosperm as well (Fig. 3-10). In agreement with our failure to detect upregulation of *LEC2* in *mea/MEA* mutants, we did not observe any differences in intensity and distribution of *LEC2::GUS* expression between wild-type and the mutant seeds (data not shown).

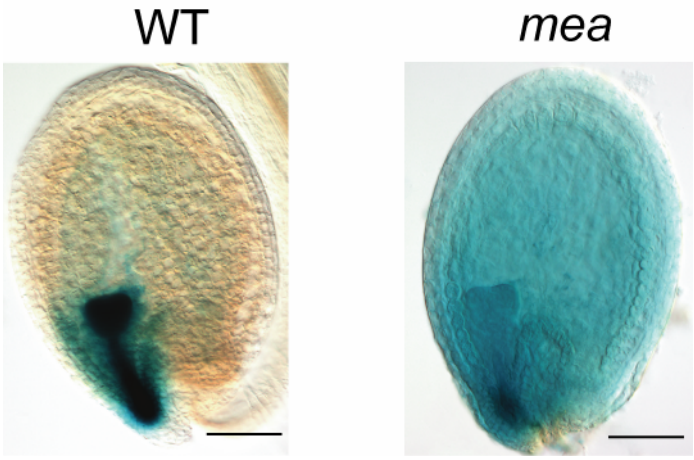


Fig 3-10 | *FUS3* is overexpressed in the endosperm. FUS3::GUS expression in wt and *mea* mutant seeds. Scale bars, 100 µm.

We analyzed whether *FUS3* is directly targeted by MEA and marked by H3K27me3 trimethylation by performing ChIP experiments with anti-MEA and anti-H3K27me3 antibodies. With both antibodies we detected significant enrichment of the *FUS3* promoter locus in chromatin from closed wild-type flowers, establishing *FUS3* as a direct target gene of MEA.

Taken together, these data are in favor of the hypothesis that there are at least several genes which are regulated by different PRC2-like complexes at different developmental stages.

3.2 Targeting by FIS-complex followed by trimethylation of H3K27 is required but not sufficient for genomic imprinting of *PHE1*

3.2.1 *PHERES2* is a direct target gene of FIS-complex

Analysis of the genome sequence of *Arabidopsis* revealed the presence of a close *PHE1* homolog that we named *PHERES2 (PHE2). PHE2* is localized on the same chromosome in antisense orientation approx. 10 kbps apart from *PHE1*. Phylogenetic analysis of the type I MADS box gene family predicts these two genes having recently duplicated from a

common ancestor (Parenicova et al., 2003). Alignment of genomic sequences of both genes including upstream and downstream sequences revealed a high level of similarity within the coding sequences. While there is also significant similarity of approx. 450 bp of upstream sequences, sequences located downstream of stop-codons do not demonstrate significant similarity (Fig 3-11).

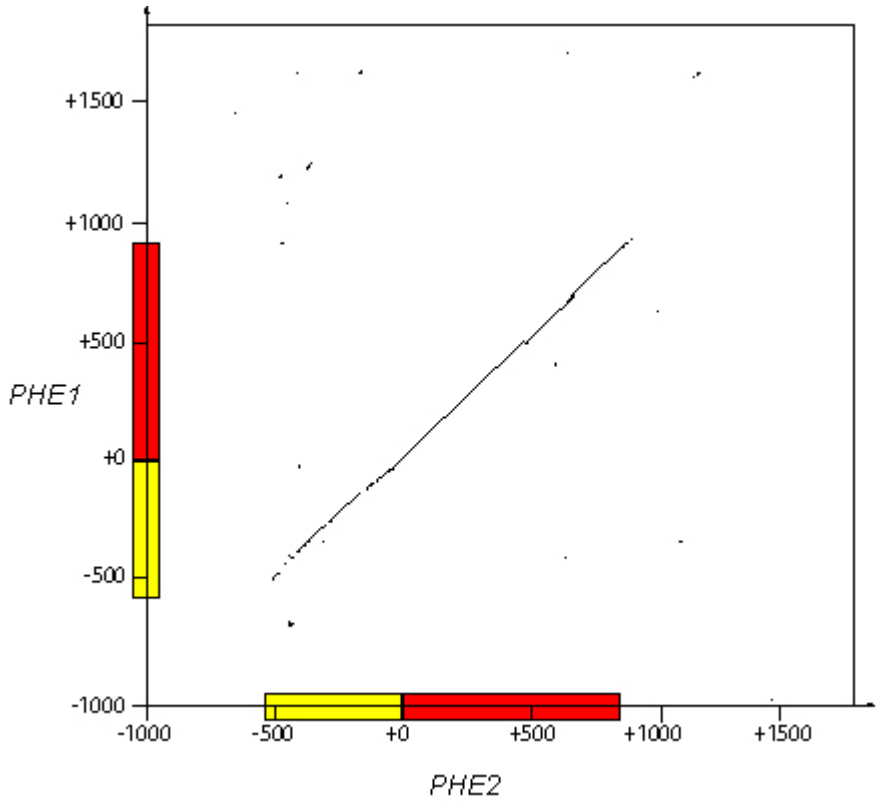


Fig 3-11 | Dot plot of sequence similarity between *PHE1* and *PHE2* loci. Numbers on the scales are coordinates relative start codons (+0) of the genes. Regions corresponding to ORFs are highlighted with red color. Upstream sequences sharing high similarity are highlighted with yellow color.

Given the high sequence similarity in the promoter region of *PHE1* and *PHE2*, we addressed the question whether *PHE2* is also a targeted and repressed by the FIS PcG complex. We developed a PCR assay based on a deletion in the *PHE2* sequence that allowed distinguishing between *PHE1* and *PHE2* PCR products. This assay provided the

technical opportunity to analyze the results of ChIP experiments for corresponding regions of *PHE1* and *PHE2* on the same semi-quantitative electrophoresis gels. In ChIP experiments performed with chromatin from closed flowers using anti-MEA and anti-H3K27me3 antibodies we observed similar enrichment levels for promoter regions of both *PHE1* and *PHE2* (Fig 3-12).

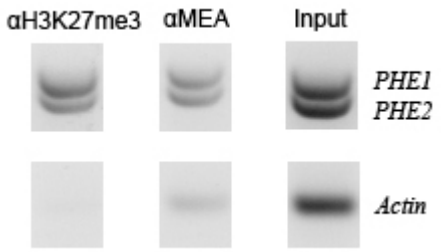


Fig 3-12 | Analysis of H3K27me3 and MEA enrichment at *PHE1* and *PHE2* loci. ChIP was performed with closed flowers of wild-type plants.

These results demonstrate that *PHE2* is also a direct target gene of MEA. Furthermore, they suggest that 450 bp promoter sequences are likely to contain all necessary information for directing MEA to these regions. To further test this prediction, we performed ChIP experiments analyzing binding of MEA and presence of H3K27me3 mark over the junction of *PHE1*promoter::*GUS* in the transgenic lines DB203 and DB215, expressing the *GUS* reporter gene under control of the 800bp *PHE1* promoter. We observed that MEA binding to this artificial construct was comparable to MEA binding to the endogenous locus. Enrichment levels of H3K27me3 at the *PHE1*promoter::*GUS* junction region was significantly lower compared to the endogenous locus, however, still detectable (Fig. 3-13).

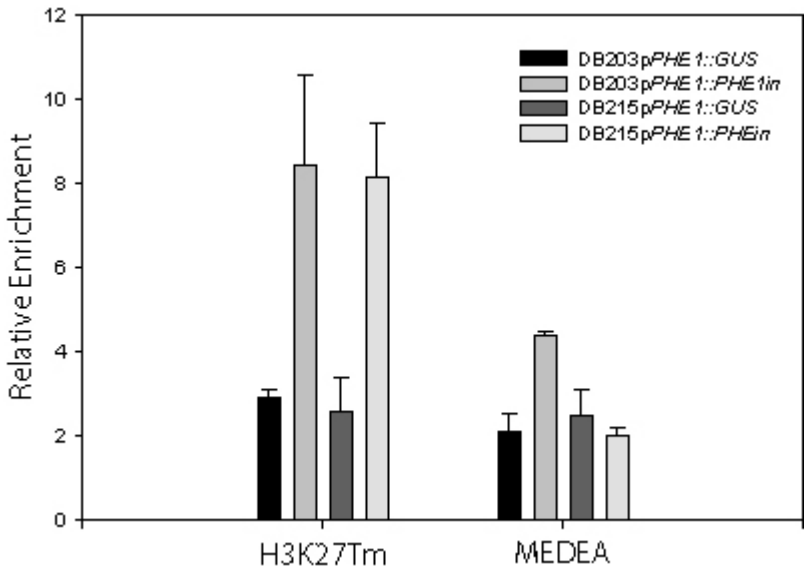


Fig 3-13 | 800bp *PHE1* promoter contains sequence context required for MEA binding and to establish H3K27me3. ChIP samples with H3K27me3 and MEA antibodies were analyzed by realtime PCR in triplicate. Material used: closed flowers of transgenic lines DB203 and DB215, containing a *GUS* gene under control of 800bp of *PHE1* promoter. Regions analyzed: pPHE::GUS – junction of 800bp *PHE1* promoter with *GUS* gene; p*PHE::PHEin* – region 2 (Fig. 3-1).

3.2.2 *PHE2* is not upregulated in the *mea* mutant

To further investigate the impact of MEA on regulation of *PHERES2* expression, we asked whether *PHE2* is upregulated in the *mea* mutant background. To address this question we measured relative *PHE2* expression levels in *mea* mutant plants pollinated with wild-type pollen in siliques harvested at 1, 2, 3, and 4 DAP (Fig. 3-14). Surprisingly, in contrast to *PHE1* we failed to detect upregulation of *PHE2* expression in *mea* mutant background. This result indicates that additional regulators impact on the regulation of *PHE1* and *PHE2* expression.

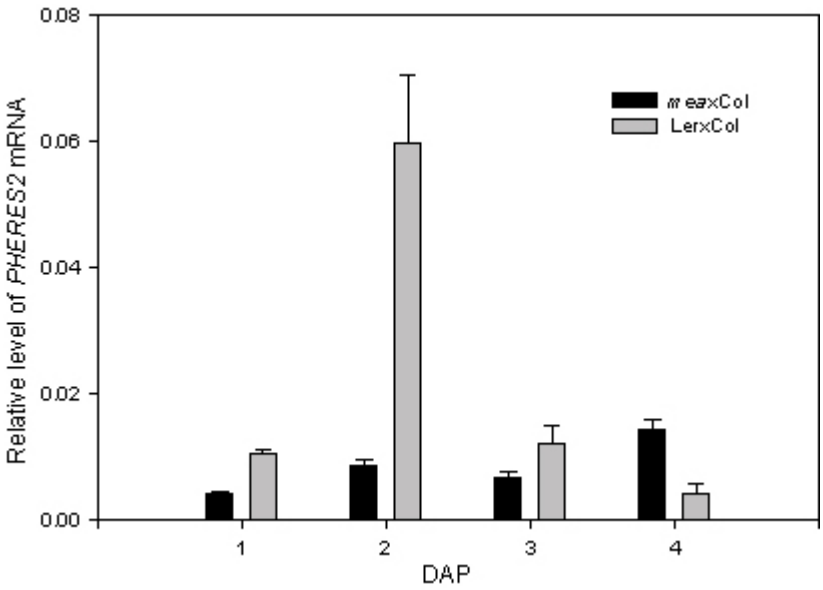


Fig 3-14 | *PHE2* is not upregulated in *mea* mutants. Relative *PHE2* mRNA levels in siliques harvested at 1-4 DAP resulting from crosses of *mea* mutants or L*er* plants with Col plants.

3.2.3 *PHE2* is not controlled by genomic imprinting

PHE1 is regulated by genomic imprinting, resulting in a preferential expression of the paternal *PHE1* allele (Kohler et al., 2005). Having identified *PHE2* as a direct target gene of MEA, we addressed the question whether genomic imprinting is a common trait of all MEA regulated target genes. We made use of a single nucleotide polymorphism (SNP) in the *PHE2* sequence between the accessions *Columbia*(Col) and Landsberg *erecta* (L*er*), which could be traced using the restriction endonuclease TaqI. We assessed relative expression levels of maternally and paternally derived *PHE2* alleles in siliques at 1-4 DAP obtained from crosses of L*er* and Col accessions. In contrast to the preferential paternal expression of *PHE1*, *PHE2* was clearly biallelically expressed (Fig. 3-15).

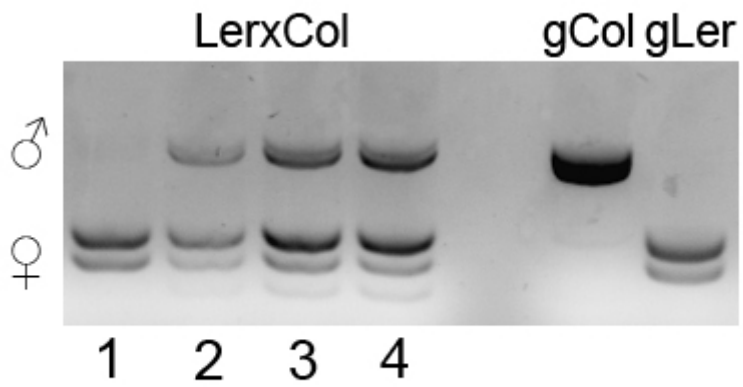


Fig 3-15 | *PHE2* is biallelically expressed at early developmental stages. Allelic expression of *PHE2* was assessed in siliques resulting from pollination of L*er* plants with pollen of Col plants. RT-PCR products were digested with *Taq*I. Genomic DNA of Col (gCol) and Ler (gLer) is shown as control.

Thus, although *PHE2* is targeted by MEA and contains H3K27me3 marks at the promoter region, it is not regulated by genomic imprinting. This strongly indicates that MEA binding followed by trimethylation of H3K27 is required but not sufficient for genomic imprinting at the *PHE1* locus.

3.3 Analysis of imprinting status of *PHE1* in mutants affecting different gene silencing pathways

3.3.1 Experimental approach

Our previous results (Makarevich et al., 2006) demonstrated that inactivation of the maternal *PHE1* allele is associated with chromatin modifications, mainly H3K27me3. However, our initial experiments were not designed to address the question by which mechanisms the FIS complex is specifically targeted to the maternal *PHE1* allele. To address this question, we analyzed the parent of origin dependent expression of *PHE1* in several mutants previously shown to be impaired in establishing and/or maintenance of

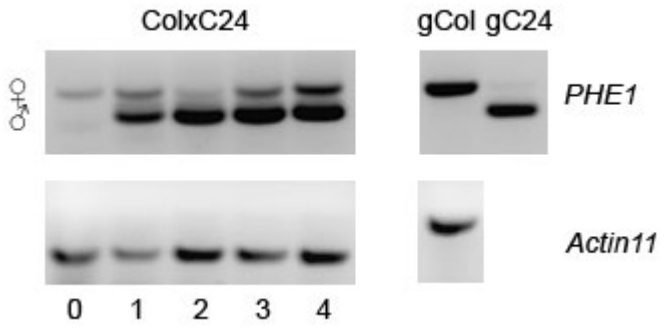
gene silencing..

The mutants chosen to be analyzed are either in the *Columbia(Col)* or *Wassilewskija(Ws)* background. This allowed us to use the established assay for testing allele specific *PHE1* expression by exploiting the presence of a SNP between these accessions and the *C24* accession (Kohler et al., 2005).

3.3.2 *PHE1* imprinting in mutants affecting components of siRNA pathways – *rdr2*, *dcl3*

Based on the previously reported cases of involvement of siRNA machinery in directed silencing of endogenous genomic loci (Chan et al., 2004), we analyzed whether a similar pathway participates in the establishing and maintenance of the inactive state of the maternal *PHE1* allele. *S*ilencing of a transgenic *FWA* locus by DRM2 mediated *de novo* DNA methylation requires components of the siRNA machinery, in particular *DICER-LIKE 3* (*DCL3*) and *RNA POLYMERASE 2* (*RDR2*) (Chan et al., 2004). To test if imprinting of the maternal *PHE1* allele also depends on the siRNA machinery, we analyzed the parent of origin dependent expression of *PHE1* in *dcl3* and *rdr2* mutants. We measured maternally and paternally derived *PHE1* transcripts in the crosses between these mutants (which are in *Columbia* ecotype) used as maternal plants and *C24* wild-type plants pollen donors. In both cases we did not observe any differences in expression compared to control experiments where wild type *Col* plants were used instead of mutants (Fig.3-16).

A



B

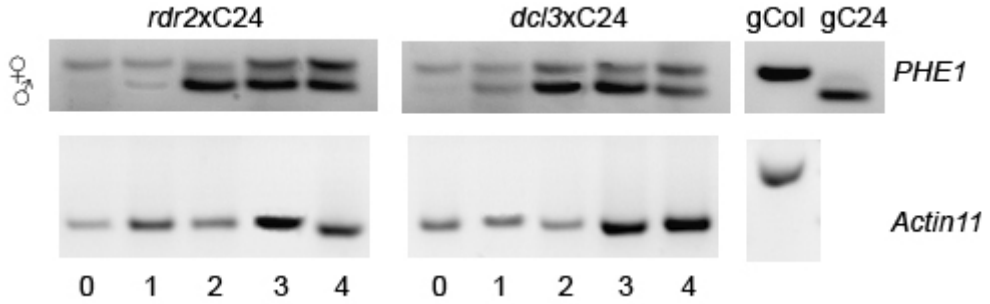


Fig 3-16 | Imprinting of maternal allele of *PHE1* is not affected in *rdr2* and *dcl3* mutants. (A) Wild-type *Col* plants and (B) *rdr2* and *dcl3* mutants (*Col*) were crossed with wild-type *C24* plants and allelic expression of *PHE1* was analyzed using a restriction site polymorphism between maternal and paternal alleles.

3.3.3 *PHE1* imprinting in mutants affecting DNA-methylation – *drm1/drm2*, *met1*, *cmt3*, *kyp*

DNA methylation is one of two principal mechanisms regulating gene expression. Previous reports on imprinting in *Arabidopsis*, demonstrated the importance of DNA methylation in imprinting regulation of *MEA* and *FWA* (Gehring et al., 2006; Kinoshita et al., 2004). To analyze whether the DNA methylation machinery is important for the imprinting control of the *PHE1* locus, we analyzed parental expression of *PHE1* in mutants with affected components of the DNA *de-novo* methylation pathway and components of the DNA maintenance methylation machinery. This analysis was

performed using the same experimental design as described above for the analysis of *PHE1* expression in *dcl3* and *rdr2* mutants.

DRM1 and DRM2 proteins represent *de novo* DNA methyltransferases in *Arabidopsis*. Double *drm1/drm2* mutants are unable to efficiently methylate transgenes containing *FWA* and/or *SUP* loci (Cao and Jacobsen, 2002). At the same time, DRM1, DRM2 are not required for propagation of previously established DNA methylation at the same loci (Cao and Jacobsen, 2002). We tested whether *de novo* DNA methylation of *PHE1* in the central cell is required for silencing of the maternal allele by analyzing parental expression of *PHE1* after crossing *drm1/drm2* mutants with wild-type C24 pollen. As shown in Fig 3-17, in the products of such cross we observed preferential expression of the paternal *PHE1* allele, indicating that *de novo* DNA methylation is not required for silencing of the maternal *PHE1* allele.

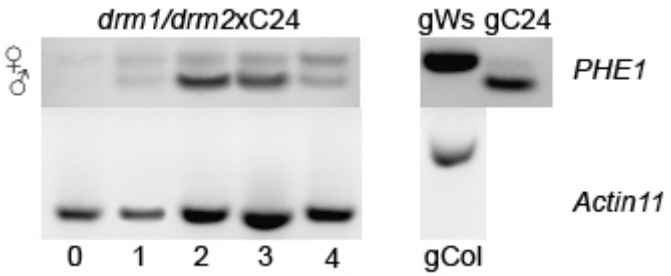


Fig 3-17 | Imprinting of the maternal allele of *PHE1* is not affected in *drm1/drm2* double mutants. *drm1/drm2* mutants (*Ws*) were crossed with with wild type *C24* plants and allelic expression of *PHE1* was analyzed using a restriction site polymorphism between maternal and paternal alleles.

Mutants in the *MET1* gene encoding a DNA methyltransferase demonstrated global loss of DNA methylation and reactivation of silenced transgenes (Kankel et al., 2003; Saze et al., 2003). MET1 is required for maintenance of DNA methylation, particularly in the hemimethylated CpG context during DNA replication. Although *de novo* methylation did not impact on maternal *PHE1* silencing, we considered the possibility that *PHE1* is regulated by DNA methylation that is not reset in each generation. To test this

hypothesis, we analyzed the parental pattern of *PHE1* expression in crosses of *met1* mutants with wild-type *C24* plants. As it is shown on Fig 3-18, there is no reactivation of the maternal *PHE1* allele in the products of these crosses, indicating that MET1 is not required for maintaining the inactive state of the maternal allele.

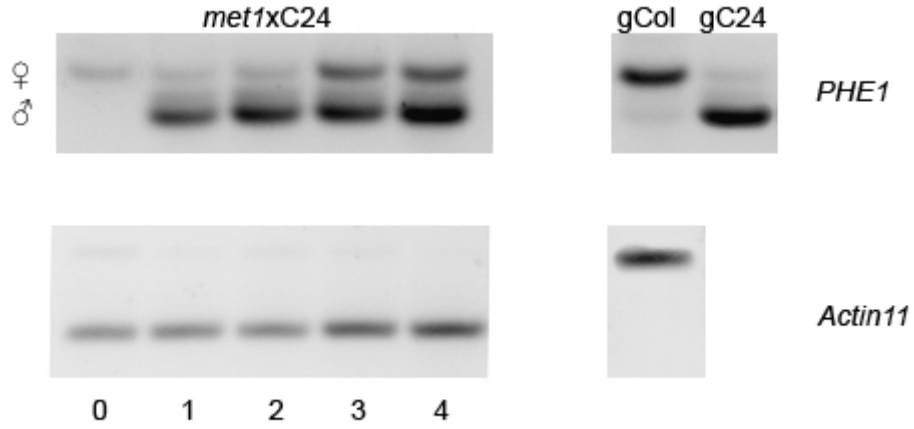


Fig 3-18 | Imprinting of the maternal allele of *PHE1* is not affected in *met1* mutants. *met1* mutants (*Col*) were crossed with wild type *C24* plants and allelic expression of *PHE1* was analyzed using a restriction site polymorphism between maternal and paternal alleles.

In parallel with the analysis of *met1* mutants, we examined *PHE1* expression in *cmt3* and *kyp* mutants that are impaired in maintaining methylation at asymmetric and CpNpG sites (Bartee et al., 2001; Lindroth et al., 2001; Lindroth et al., 2004; Malagnac et al., 2002). Using the same experimental setup, we analyzed parental expression of *PHE1* in crosses of the mutants (*Col*) with wild-type *C24* plants. As shown in Fig. 3-19 we did not observe maternal allele reactivation in these crosses.

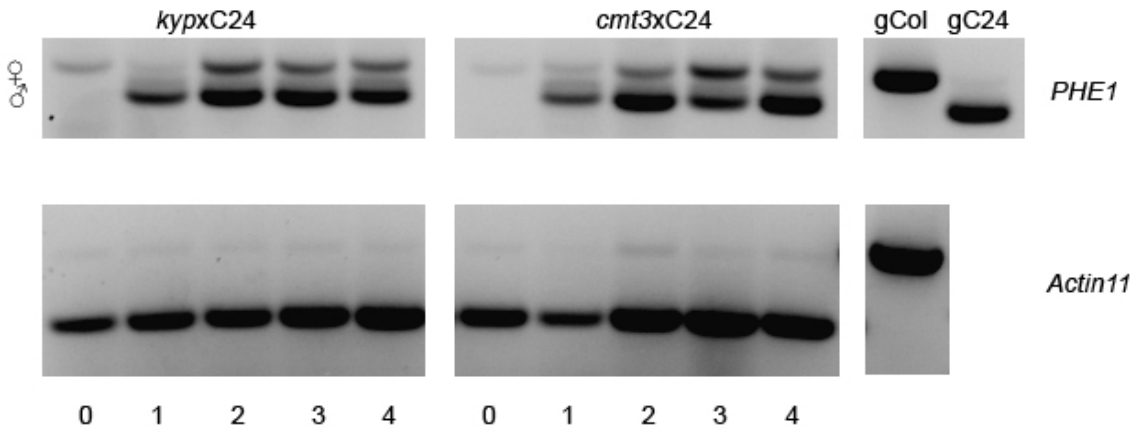


Fig 3-19 | Imprinting of the maternal allele of *PHE1* is not affected in *cmt3* and *kyp* mutants. *cmt3* and *kyp* mutants (*Col*) were crossed with wild type *C24* plants and allelic expression of *PHE1* was analyzed using a restriction site polymorphism between maternal and paternal alleles.

Redundancy in DNA methylation pathways does not allow to completely exclude a possible role of DNA methylation in silencing of the maternal *PHE1* allele. However, altogether, our results do not support a role of DNA methylation in the imprinted regulation of the *PHE1* maternal allele.

Despite the common assumption that increased DNA-methylation is associated with gene silencing, imprinting control at the insulin-like growth factor 2 (*Igf2*) loci in mouse is a clear exception from that rule. Activation of the paternal *Igf-2* allele depends on the methylated state of an insulator element located *in-cis* between *Igf-2* and its potential enhancer. The insulator on the maternally derived chromosome is unmethylated, blocking the activating potential of the enhancer (reviewed in Wolffe, 2000). Applying this model to *PHE1* case, it is possible that expression of the paternal *PHE1* allele depends on DNA methylation. To examine this possibility, we tested whether loss of DNA methylation in the male gametes causes decreased expression of the paternally derived *PHE1* locus., We analyzed the relative level of *PHE1* mRNA in crosses of wild-type plants pollinated either with pollen from *met1* mutants or wild-type plants. Graphs shown in Fig 3-20

clearly demonstrate that the hypomethylated status is associated with a lower expression level of the paternally derived *PHE1* gene.

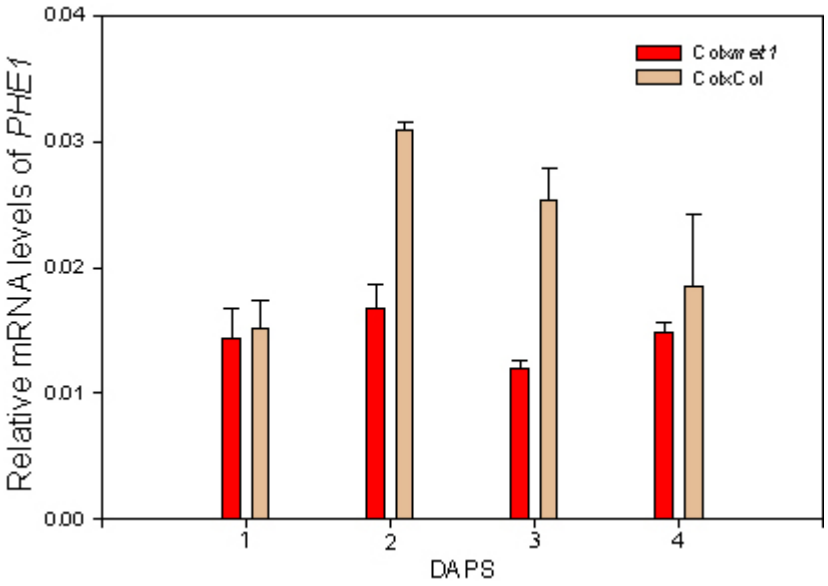


Fig 3-20 | Hypomethylation decreases expression of the paternal allele of *PHE1*. Relative level of *PHE1* mRNA measured by quantitative RT-PCR in crosses of wild-type plants pollinated either with pollen of *met1* mutants or wild-type plants.

This conclusion is additionally supported by the previously reported decrease of *PHE1* expression in *ddm1* mutants that are characterized by a significant decrease in DNA methylation. Furthermore, *PHE1* upregulation is much reduced in double *ddm1/mea* mutants compared to *mea* single mutants (Kohler et al., 2003b). However, a direct interpretation of above mentioned results is complicated by significant developmental deviations in *ddm1* mutants compared to wild-type plants.

To analyze a possible functional linkage between DNA methylation and chromatin modifications at the *PHE1* locus, we performed chromatin immunoprecipitation assays with cross-linked chromatin from closed flowers with antibodies against H3K27me3 and MEA in *met1* mutants. As shown on Fig. 3-21, we observed a significant decrease in enrichment levels for MEA as well as H3K27me3 at the promoter region of *PHE1* in

met1 mutants.

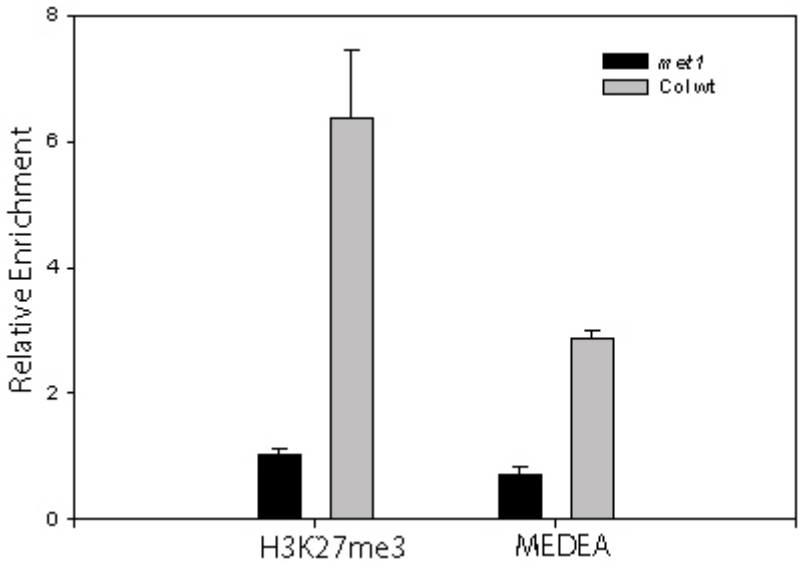


Fig 3-21 | Genome hypomethylation is associated with loss of MEA binding and H3K27me3 at the *PHE1* locus. ChIP analysis of wild-type (wt), and *met1* mutant flowers was performed with anti-MEA and anti-H3K27me3 antibodies and enrichment for region 2 of *PHE1* was analyzed by realtime PCR in triplicate.

Altogether, these observations argue for the participation of a DNA methylation dependent molecular mechanisms in the imprinting regulation of *PHE1* gene; however, the question about the functional characteristics of this involvement is left open.

3.4 Role of 3' downstream region in the regulation of *PHE1* imprinting

Detailed analysis of FIS-complex binding and chromatin modification patterns in the promoter regions of *PHE1* and *PHE2*, accompanied by investigations of the imprinting status of both genes, clearly demonstrated that binding of the FIS-complex to promoter regions followed by chromatin modifications is required but not sufficient for establishment and/or maintaining of imprinting. This conclusion is supported by the analysis of FIS-binding and H3K27me3 in transgenic lines containing only promoter regions of *PHE1* fused to a GUS-reporter gene. The promoter of *PHE1* in these constructs is bound by the FIS-complex and is significantly enriched for H3K27me3

marks (Fig.3-13 and data not shown). However, these constructs still demonstrate significant maternal expression of the GUS-reporter when maternally inherited.

These results indicate that besides the upstream promoter, other genomic regions could be important for establishment of *PHE1* imprinting. As mentioned before, in contrast to coding parts and approx. 450bp of upstream sequences, the sequences located downstream of the stop-codons of *PHE1* and *PHE2* do not have significant similarity. Therefore, we proposed that the difference in the imprinting status of *PHE1* and *PHE2* is determined by the presence of a hypothetical imprinting center in the sequences located downstream of the *PHE1* stop-codon.

3.4.1 Analysis of the *PHE1* imprinting status in the line SALK_023774

To address the question whether the genomic region located after the *PHE1* stop-codon is important for silencing of the maternal *PHE1* allele, we made use of a T-DNA line carrying a T-DNA insertion approximately 700bp after stop-codon of *PHE1* (SALK_023774). We analyzed the parental *PHE1* expression pattern in reciprocal crosses of homozygous mutant plants with C24 wild-type plants. We did not observe changes in the imprinting status of *PHE1* when the T-DNA insertion allele was inherited from the paternal side (Fig. 3-22, B). However, when the same allele was inherited from the maternal side, we observed preferential maternal expression of *PHE1* (Fig. 3-22, C).

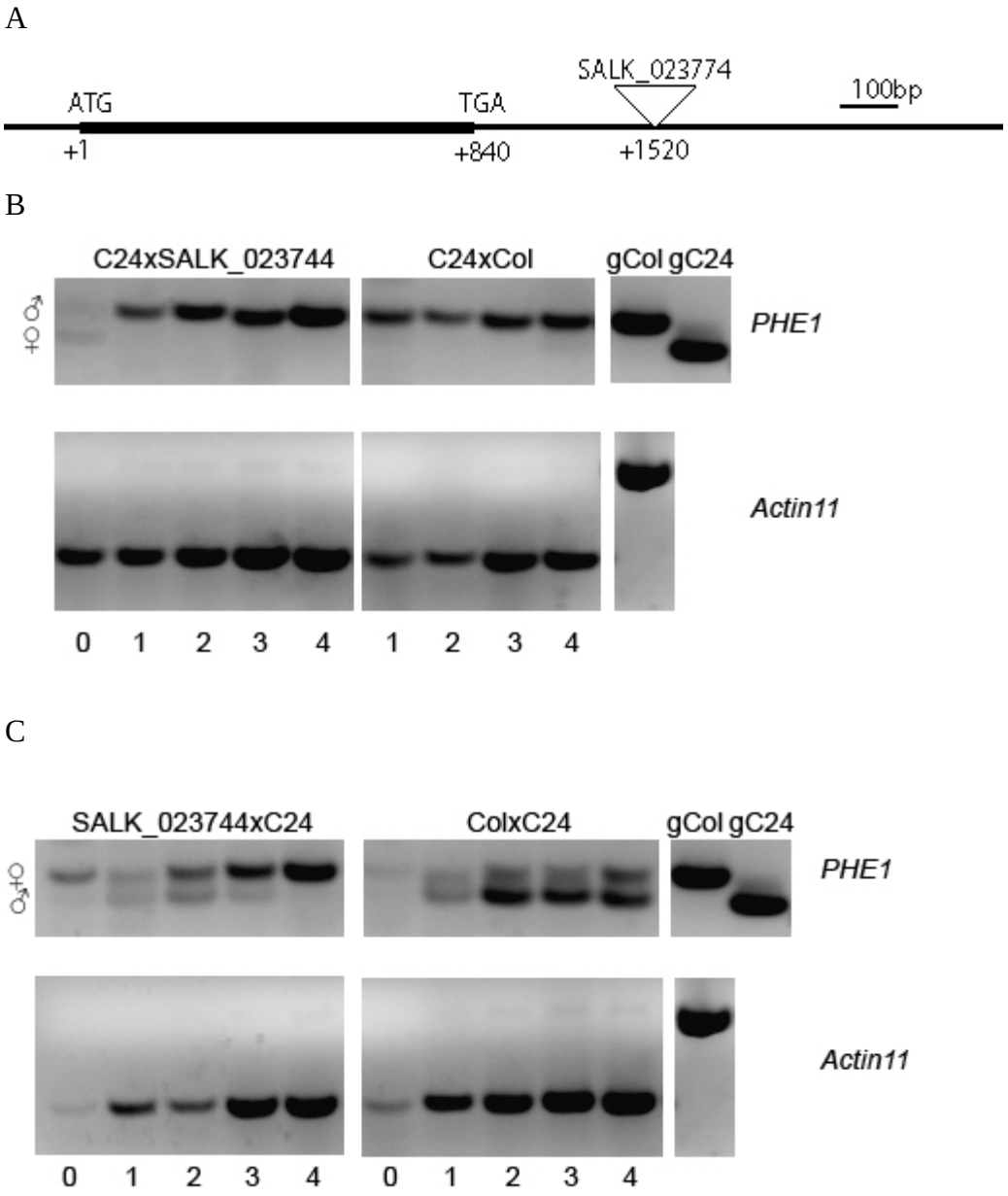


Fig 3-22 | Analysis of the parental *PHE1* expression pattern in reciprocal crosses of SALK_023774 with wild-type C24 plants. (A) Schematic diagram of the *PHE1* locus indicating the T-DNA insertion site in SALK_023744. (B) Analysis of imprinting status of *PHE1* expression in the cross C24 x SALK_023744. (C) Analysis of imprinting status of *PHE1* expression in the cross SALK_023744 x C24.

To exclude the possibility that upregulation of the maternal allele of *PHE1* in

SALK_023774 is caused by T-DNA derived promoter activity, we performed northern-blot analysis of RNA from early seedlings of SALK_023774 line using probes flanking the left and right T-DNA insertion sites. We detected a clear signal for our kanamycin control probe, however, no signal was detected with the probes for the left and right flanking sequences (Fig 3-23). Thus, we concluded that reactivation of the maternal *PHERES1PHE1* allele in SALK_023774 is caused by loss of functional control between *PHE1* and its regulating regions by insertion of the T-DNA.

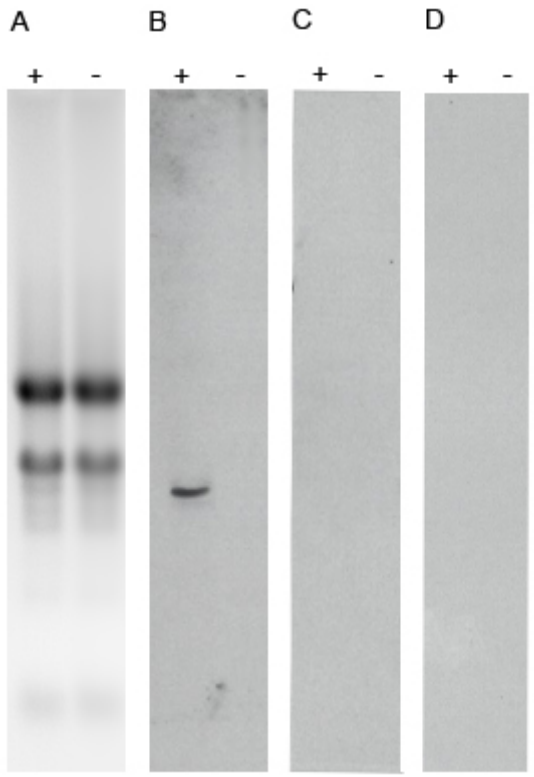


Fig 3-23 | Nothern-blot analysis of PHE1 region in the line SALK_023744. (A) Image of the gel before the transfer (B) Scan of the hybridization with the probe against kanamycine (C) Scan of the hybridization with the probe flanking the left side of the T-DNA insertion site (Fig. 3-22) (D) Scan of the hybridization with the probe flanking the right side of the T-DNA insertion site (Fig. 3-22). (+) - RNA extracted from seedlings of the SALK_023774 line; (-) RNA extracted from seedlings of *Col* wt.

To analyze whether the T-DNA insertion in SALK_023774 also reactivates *PHE1* expression in tissues where *PHE1* is normally not expressed, we performed quantitative RT-PCR analysis of RNA extracted from leaves of this line. As shown in Fig. 3-24, *PHE1* expression was indeed weakly upregulated in leaves of SALK_023774; however, the level of *PHE1* expression in this line was significantly lower compared to *PHE1* expression in leaves of *pkl* mutants.

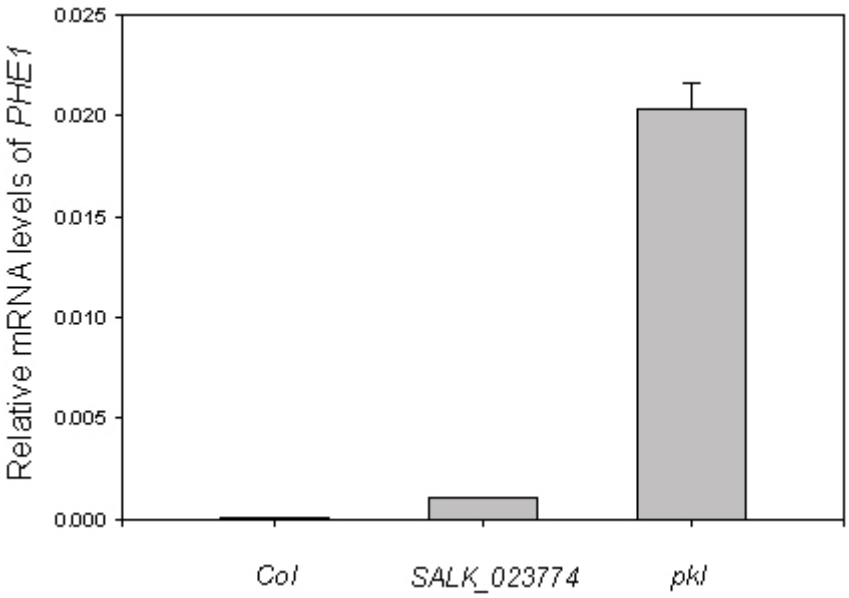


Fig 3-24 | *PHE1* expression is weakly upregulated in leaves of SALK_023774. Relative *PHE1* mRNA levels were analyzed in leaves of wt, SALK_023774 and *pkl* mutants.

3.4.2 Analysis of the DNA-methylation status in *PHE1* downstream regions

Previous attempts failed to identify DNA methylation sites in promoter as well as in coding regions of *PHE1* (Claudia Kohler, unpublished data). Our initial results suggesting a DNA-methylation dependent mechanism in the regulation of *PHE1* imprinting have been summarized in chapter 3.3.3 of this thesis. After having obtained evidence for the importance of the *PHE1* downstream regions, we hypothesized that the *PHE1* downstream regions contain DNA methylation marks involved in imprinting control of *PHE1*. To test this hypothesis, we performed methylation sensitive Southern-

blot analysis of DNA from closed flowers of *wt* and *met1* mutant plants. Using the DNA-methylation sensitive *Cla*I restriction enzyme, we could monitor the methylation status of a region located downstream of a triple direct repeat. As shown in Fig. 3-25, this site was not cut in DNA of wt flowers. In contrast, the same site was cut in DNA of *met1* mutant flowers.

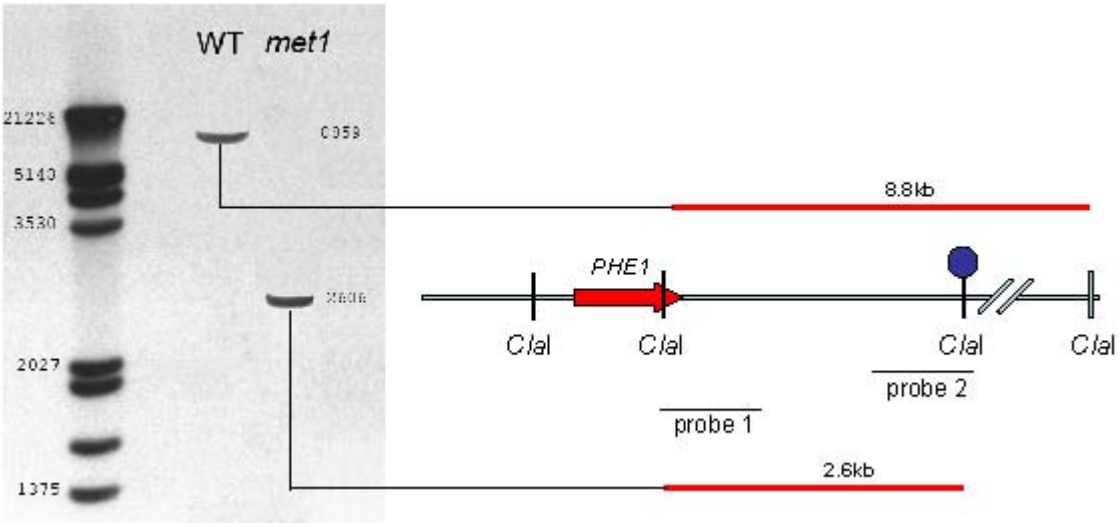


Fig 3-25 | The 3' region of *PHE1* contains DNA methylation marks. Methylation of DNA from closed flowers of wt Col plants was analyzed on southern-blots hybridized with the probes indicated. The *ClaI* restriction site that is methylated in wt and unmethylated in *met1* is indicated by a blue circle. Probes 1 and 2 (described in materials and methods as *PHE1*.downstream.1 and *PHE1*.downstream.2 respectively) are indicated.

If DNA methylation in this region is important for genomic imprinting of *PHE1*, differences in the DNA methylation frequency between paternal and maternal *PHE1* alleles would be expected. However, for technical reasons a direct comparison of DNA methylation between pollen and female gametophytes was not possible. In order to understand whether there are DNA methylation differences between gametophytic and sporophytic tissues, we performed bisulphite sequencing of DNA from pollen, closed flowers, leaves, old siliques, and early seedlings of *Col* wt plants. We analyzed the region surrounding the *Cla*I site that was found to be methylated in flowers. In the analyzed region we identified 4 sites of CpG DNA methylation and 1 site of CpTpG DNA

methylation, including the site, predicted based on southern-blot analysis (Fig. 3-26).

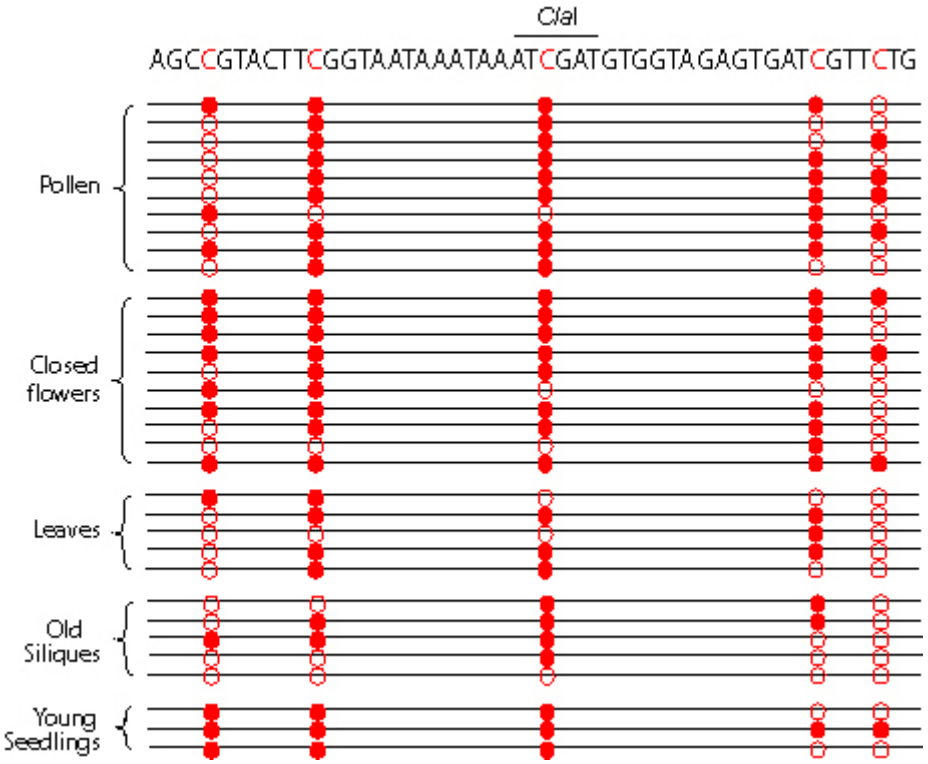


Fig 3-26 | Distribution of DNA methylation analyzed by bisulphite sequencing at the region surrounding the *Cla*I site, which was found to be methylated in flowers. Methylated cytosines are highlighted with filled red circles, unmethylated cytosines are identified by open circles. Lines represent sequenced clones.

Our analysis did not reveal a general tendency of lower DNA methylation at the identified sites in sporophytic tissues compared to pollen. In agreement with our findings that in *met1* mutants paternal *PHE1* expression is reduced, we hypothesize that DNA methylation is associated with the active *PHE1* locus. Consequently, if DNA methylation at these sites is important for *PHE1* imprinting, we need to assume that these sites become demethylated in the central cell of the female gametophyte. As the endosperm is not inherited, this prediction would exclude the need of special remethylation mechanisms of the maternally derived *PHE1* allele in sporophytic tissues.

3.5 Search for antisense RNA regulating *PHE1* expression

Investigations of genomic imprinting in animal model systems report the existence of at least one noncoding RNA for every locus studied so far (reviewed in Pauler and Barlow, 2006). Investigations of genomic imprinting in *Kcnq1*-, *Igf2r*-, *Pws*-, and *Gnas*-imprinted clusters in mice revealed transcription of nonconding RNA in antisense orientation to at least one of the genes in the imprinted cluster. Therefore, we analyzed whether antisense oriented noncoding RNAs participate in *PHE1* imprinting. For this purpose we performed reverse transcription with primers designed specifically to initiate first-strand cDNA synthesis from a hypothetical RNA followed by a PCR assay. We included the middle part of the *PHE1* coding region and the region overlapping the 3' end of *PHE1* in our analysis. (Fig. 3-27). As positive control for region 1 we performed the same assay but using the antisense primer for first strand cDNA synthesis. As positive control for region 2 we performed a similar assay using oligo-dT primers for first strand synthesis.

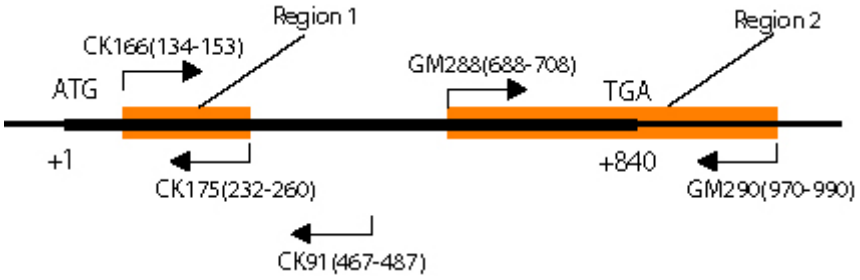
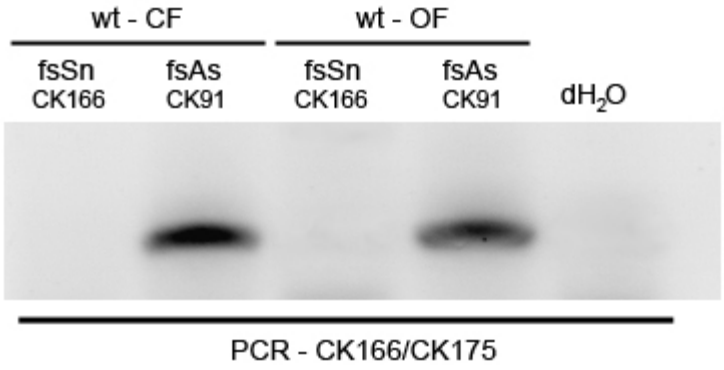


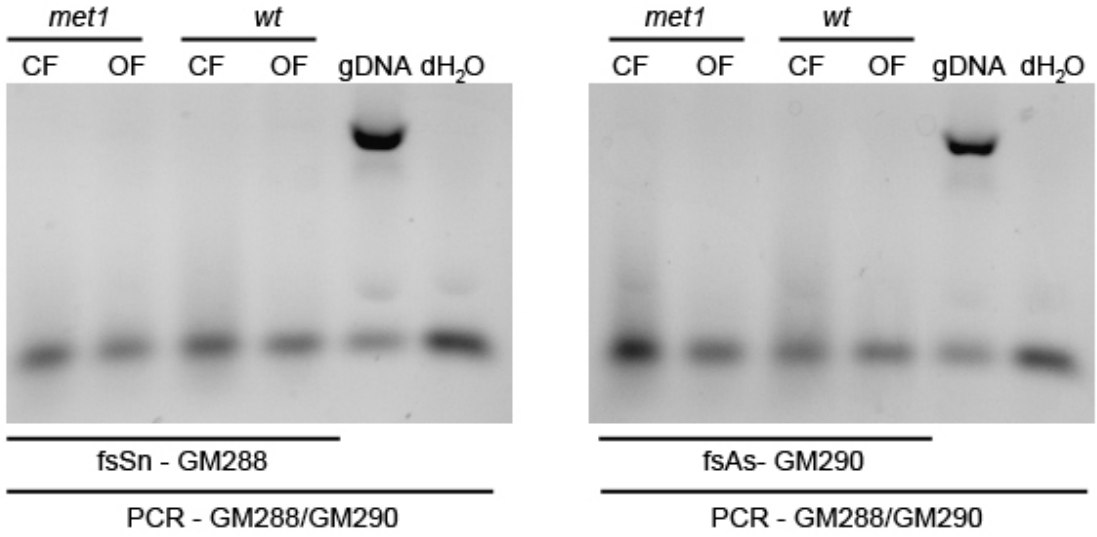
Fig 3-27 | Schematic presentation of regions 1 and 2of the *PHE1* locus analyzed for the presence of antisense non-coding RNA(s).

We tested RNA from closed and open flowers of *Col* wt plants, but as demonstrated in Fig. 3-28, we failed to detect antisense transcripts originating from the analyzed regions.

A



B



C

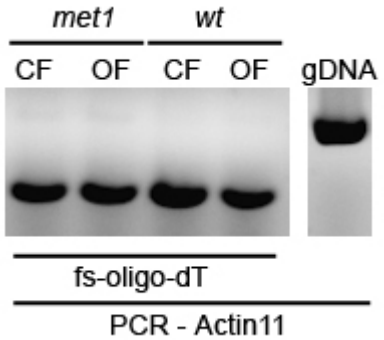


Fig 3-28 | The *PHE1* locus does not form noncoding RNAs. RT-PCR analysis was performed with RNA from closed (CF) and open (OF) flowers of *Col* wt and *met1* mutant

plants. First strand synthesis in reverse transcription was performed using one of sense (fsSn), antisense (fsAs), or oligo-dT primers (fs-oligo-dT). The products of first strand synthesis reaction were analyzed by PCR with primers shown in Fig. 3-27. (A) RT-PCR analysis of region 1 (Fig. 3-27) in *Col* wt plants. The reaction with antisense primer fsAs was used as positive control. (B) RT-PCR analysis of region 2 (Fig. 3-27) in *Col* wt and *met1* mutant plants. The reaction with antisense primer fsAs was used as negative control. (C) Positive control for the reactions shown in (B). First strand synthesis was performed with oligo-dT primers. PCR was performed with primers specific to *Actin11*.

The failure to identify potential antisense oriented RNA transcripts in regions overlapping the transcribed *PHE1* gene combined with previously described failure to observe changes in the imprinting status of *PHE1* in *dcl3* and *rdr2* mutants impaired in components of the RNAi machinery argues against the hypothesis that antisense RNA transcripts are involved in regulation of *PHE1* imprinting. However, it remains possible that noncoding RNAs are formed in regions that were not included in our analysis.

4 Discussion

Epigenetic phenomena combining changes in the gene expression status that are not associated with direct modifications of the primary coding context is a hot topic of current molecular genetics research. Many fundamental biological and biomedical processes, including cell differentiation in multicellular organisms and, in many cases, cancer formation are regulated by epigenetic mechanisms. Therefore, understanding of the molecular mechanisms causing epigenetic modifications is important for further progress in both fundamental biology and development of medical applications.

Genomic imprinting, a mechanism in which the expression status of an allele depends on its parental origin, is one of most prominent examples of epigenetic phenomena. It is also associated with unusual genetic inheritance of several human genetic diseases, including the well-studied Prader-Willi/Angelman (PWS/AS) and the Beckwith–Wiedemann (BWS) syndromes (reviewed in Kantor et al., 2006; Soejima and Wagstaff, 2005). It makes genomic imprinting a good model to study epigenetic phenomena. Despite a long history of research in this field, understanding of the molecular basis of this phenomenon is not yet complete.

PHE1 is the only imprinted gene known that is paternally expressed at early developmental stages in *Arabidopsis* (Kohler et al., 2005). Therefore, *PHE1* can be used as a model to study imprinting phenomena in plants. In this work I tried to investigate the role of basic epigenetic mechanisms in genomic imprinting of *PHE1* and to discover potential imprinting control regions in the *PHE1* locus.

4.1 Regulation of *PHE1* expression through chromatin modifications at the locus

Previous studies demonstrated that *PHE1* is a direct target gene of the FIS PRC2-like complex active in seed endosperm at early developmental stages (Kohler et al., 2003b). Combined with the previously demonstrated role of PcG complexes in epigenetic chromatin modifications through histone methylation in different animal models (Cao et

al., 2002; Czermin et al., 2002; Muller et al., 2002), this fact strongly argued for the hypothesis that in *Arabidopsis PHE1* is similarly regulated by the FIS-complex. Therefore, we have analyzed the distribution of several known histone H3 methylation modifications at the *PHE1* locus at different developmental stages. Indeed, we have observed significant enrichment for H3K27me3 mark over promoter region of *PHE1*. Similar enrichment for H3K27me3 over promoter regions in *Arabidopsis* was previously shown for other targets of PcG-like complexes, specifically, *FLOWERING LOCUS C* (*FLC*) (Bastow et al., 2004; Sung and Amasino, 2004) (targeted by VRN2); *MEA* (Gehring et al., 2006; Jullien et al., 2006a) (targeted by MEA and CLF); *AGAMOUS* and *SHOOTMERISTEMLESS* (*STM*) (Schubert et al., 2006) (targeted by CLF). In addition to H3K27me3, we found that the *PHE1* promoter contains H3K9me2 in closed flowers – another epigenetic mark of inactivated chromatin. In agreement with previously proposed negative control model for *PHE1* (Kohler et al., 2003b), we did not detect active chromatin epigenetic marks like H3K4me2 in any of the analyzed developmental stages.

We observed a remarkable correlation between the *MEA* expression profile, binding of MEA and enrichment of H3K27me3 at the promoter region of *PHE1*. This strongly argues for the hypothesis that H3K27me3 at the *PHE1* region in gametophytic tissues and developing seeds is established by a MEA-dependent PcG complex. This hypothesis is further supported by a detected decrease of enrichment for H3K27me3 at the *PHE1* locus in homozygous *mea* mutants.

Despite the decrease of H3K27me3 levels in *mea* homozygous mutants, the remaining enrichment level at the *PHE1* locus predicts the existence of another MEA-independent molecular mechanism able to establish this epigenetic mark. Our analysis of chromatin modifications at the *PHE1* locus in mutants lacking the *MEA* homologs – *CLF* and *SWN* – revealed that CLF and SWN redundantly establish H3K27me3 at the *PHE1* promoter in sporophytic tissues. We also observed that decrease in H3K27me3 in double *clf/swn* mutants is associated with upregulated *PHE1* gene expression. However, the relative level of *PHE1* expression in *clf/swn* double mutants was significantly lower compared to the expression in reproductive tissues where *PHE1* is normally expressed. Therefore, we conclude that besides removal of inhibitory mechanisms, additional activation signals are required to activate *PHE1* expression. Alternatively, it is possible that during sporophytic

development additional silencing pathways keep *PHE1* tightly repressed, even in the absence of PcG complexes. Indeed, high *PHE1* expression in *pkl* mutant leaves support for this view.

Based on the presented observations, we predict a model (Fig 4-1) for genomic imprinting of *PHE1* by the specific activity of the FIS complex in the female gametophyte, establishing H3K27me3 chromatin marks at the maternal *PHE1* allele. The same complex is absent in the male gametophyte, therefore the paternal allele of *PHE1* does not have inactivation marks and is preferentially activated by some still unknown activating factor.

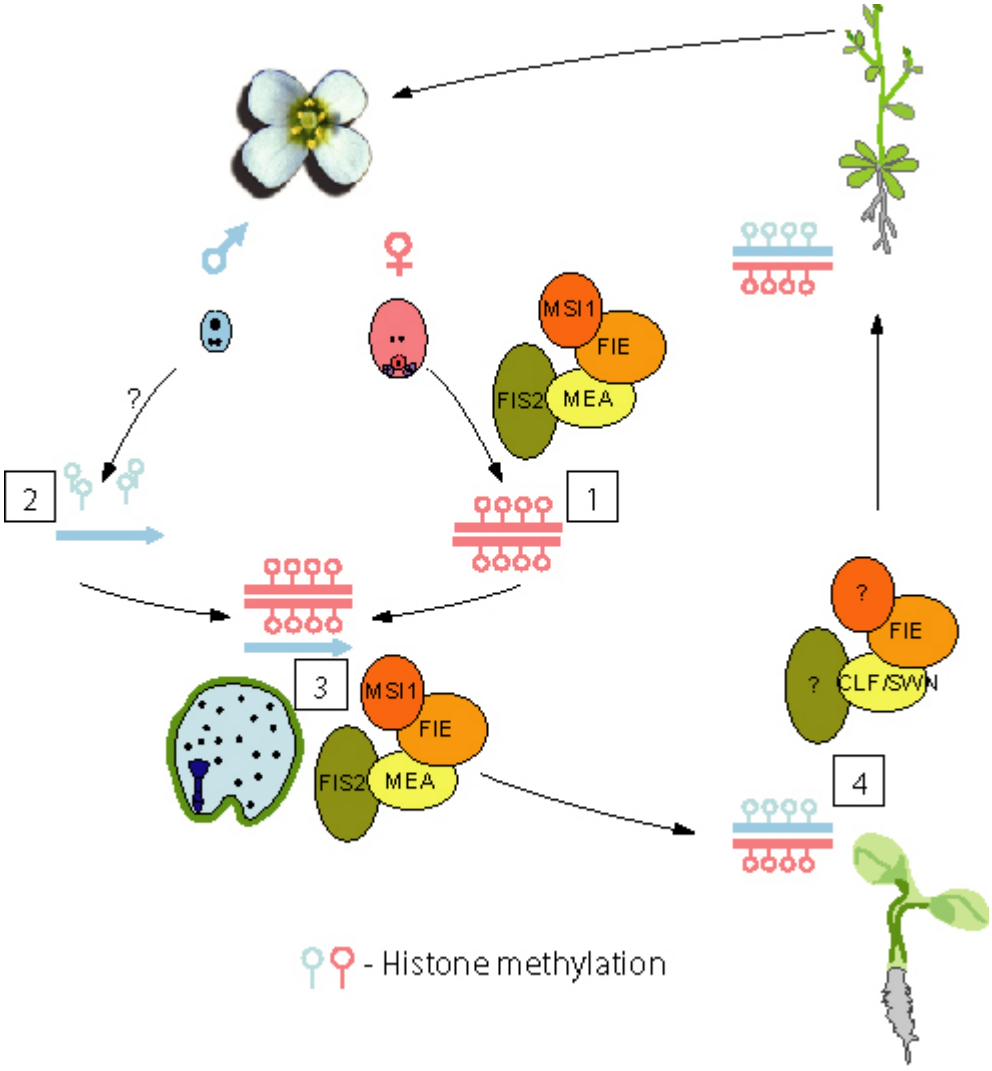


Fig 4-1 | Model of molecular mechanisms controlling genomic imprinting of the *PHE1*

locus of *Arabidopsis*. The maternally imprinted gene *PHE1* (*PHE1*; red) is repressed in the female gametophyte by the FIS PcG complex, which contains MEA, FIE, FIS2 and MSI1, and confers methylation to H3K27 1). From the active paternal *PHE1* allele (blue) H3K27 methylation is removed, but it is not known whether this is an active or passive process (2).Repression of the maternal *PHE1* allele in the endosperm is maintained by the FIS complex (3). In vegetative tissues, the silent state of *PHE1* is maintained by H3K27 methylation (4), which depends on PcG complexes containing CURLY LEAF (CLF), SWINGER (SWN) and FIE.

4.2 The FIS-complex is required but not sufficient to establish maternal imprinting of *PHE1* gene

Despite the appealing simplicity of the previously presented model, nicely predicting that H3K27me3 at the *PHE1* locus established by the maternally active FIS complex can be the primary molecular mark differentiating transcriptionally silenced maternal and active paternal alleles, our finding that the non-imprinted *PHE2* is also targeted and methylated by the FIS-complex, strongly suggests that besides the FIS-complex there are additional molecular mechanisms involved in the imprinting regulation of *PHE1*. This hypothesis is in agreement with the finding that H3K27me3 mediated by CLF is necessary but not sufficient for silencing of *AGAMOUS* (Schubert et al., 2006).

4.3 No evidence for a role of RNAi-dependent molecular pathways in *PHE1* imprinting

In all imprinted mammalian loci that have been analyzed thus far, non-coding RNAs have been detected. Depending on the predicted functional role of non-coding RNAs in imprinting establishment, two groups can be distinguished: the first group (*Kcnq1*, *Igf2r*, *Pws*, *Gnas*) includes cases, where non-coding RNAs are transcribed in antisense orientation to at least one gene of the imprinted cluster and transcription is supposed to be

important for cluster imprinting; the second group (*Igf2*, *Dlk1*) includes examples where non-coding RNAs are not supposed to be important for imprinting (reviewed in Pauler and Barlow, 2006). We hypothesized that non-coding RNAs in antisense orientation to the *PHE1* transcript could also be present at the *PHE1* locus and involved in imprinting of *PHE1*. We followed two strategies in our analysis: (i) we analyzed whether knockout of components of the RNAi-dependent gene silencing machinery affects *PHE1* imprinting, and (ii) we designed a PCR strategy that would allow us to detect antisense-oriented transcripts overlapping the *PHE1* coding region.

Our analysis of *PHE1* imprinting in *dcl3* and *rdr2* mutants failed to detect any differences in the parental expression pattern compared to the wild-type control. Previously, RDR2 and DCL3 were shown to be required for maintenance of silenced chromatin through an siRNA-directed pathway at several endogenous loci including retrotransposons, 5S rDNA, centromeric tandem repeats (Xie et al., 2004), as well as for directing *de-novo* DNA-methylation to *FWA* direct tandem repeats (Chan et al., 2004). Our failure to detect changes in the imprinting status of *PHE1* in these mutants suggests that siRNA-directed *de-novo* DNA methylation is not involved in *PHE1* imprinting. However, a high degree of functional redundancy in siRNA-dependent pathways in *Arabidopsis* (according to plant chromatin database www.chromdb.org, there are four genes encoding for Dicer-like proteins and six genes encoding for RDR-like proteins) does not allow to completely exclude this possibility and analysis of double/triple mutants among components of the siRNA pathway may be required to completely rule out this hypothesis.

We also failed to detect antisense RNA transcripts in two regions overlapping with the *PHE1* open reading frame (ORF). As the analyzed regions included both 5'- and 3'-regions of the *PHE1* ORF, it is unlikely that antisense RNAs participate in *PHE1* imprinting. However, given that antisense non-coding RNAs do not necessarily need to overlap with the gene they regulate (Sleutels et al., 2002; Zwart et al., 2001), we cannot completely rule out this hypothesis.

4.4 DNA methylation is involved in *PHE1* imprinting regulation

Following the commonly accepted view that DNA methylation is one of the two (along

with chromatin modifications) principal molecular mechanisms of epigenetic gene expression regulation, we analyzed the effect of mutations that affect DNA methylation on imprinting of the maternal *PHE1* allele. .

DRM1 and DRM2 together represent the only well-studied *de-novo* DNA methylation activity in *Arabidopsis*. Double mutants *drm1/drm2* demonstrate significant loss of *de-novo* DNA methylation in different tested genomic contexts, including transformed tandem repeats and transcribed inverted repeats (Cao et al., 2003; Cao and Jacobsen, 2002). Our failure to detect changes in the imprinting status of the maternal *PHE1* allele in *drm1*/*drm2* double mutants strongly argues against the hypothesis that differences between maternal and paternal *PHE1* alleles depend on *de-novo* DNA methylation activity.

Maintenance of DNA methylation at CpNpG contexts in *Arabidopsis* depends on the CMT3 – plant-specific DNA-methyltransferase (Lindroth et al., 2001). The hypothesis that CMT3 plays a role in *PHE1* imprinting seemed especially appealing in light of previous reports demonstrating that targeting of CMT3 to specific genomic contexts depends on histone H3 post-translational modifications, specifically simultaneous methylation of K9 and K27 (Lindroth et al., 2004). The effect on DNA methylation observed in the *cmt3* mutant is at least partially phenocopied in the *kyp* mutant, affecting the component required for H3K9-methylation (Jackson et al., 2004; Jackson et al., 2002; Malagnac et al., 2002). Therefore, we analyzed the imprinting status of the maternal *PHE1* allele in both *cmt3* and *kyp* mutants, but we did not detect any changes compared to the wild-type control.

Finally, we analyzed the effect of the *met1* mutation on maternal *PHE1* allele expression . Similarly to previously described results for *cmt3* and *kyp* mutants, we did not detect any reactivation of the maternal allele of *PHE1* in *met1* mutant background.

We also analyzed a possible functional connection between DNA-methylation, MEA binding , and H3K27 methylation at the *PHE1* promoter region by performing ChIP with the corresponding antibodies in a *met1* mutant background. Surprisingly, we observed an almost complete lack of MEA binding and H3K27me3 at the *PHE1* promoter in the *met1* mutant. Still we did not observe reactivation of the maternal *PHE1* allele. On one side, this fact clearly demonstrates that the initial model presented on Fig 4-1 is not complete,

because these results predict additional factors being necessary for repressing the maternal *PHE1* allele in the absence of H3K27me3.. This hypothesis is in agreement with our findings of only intermediate *PHE1* upregulation in *clf/swn* double mutants lacking H3K27me3 at the *PHE1* promoter. On the other side, it might indicate that DNA methylation at some yet to be identified region in the genome is required for normal reactivation of the *PHE1* gene. This explanation is consistent with our findings that *PHE1* expression is decreased in the *met1* mutant, as well as in the *ddm1* mutant (Kohler et al., 2003b), both characterized by significantly decreased overall DNA-methylation levels. Furthermore, the inability of the *mea* mutation to cause a significant upregulation of *PHE1* expression in a *ddm1* mutant background (Kohler et al., 2003b), argues in favor of this hypothesis.

4.5 The *PHE1* downstream region is required for *PHE1* imprinting

Our data indicating that the *PHE1* promoter region itself does not contain sufficient information to establish genomic imprinting of the downstream locus prompted us to analyze whether the region located downstream of *PHE1* contains a potential imprinting controlling center. To address this question, we analyzed genomic imprinting of *PHE1* in the line SALK_023774, which has approx. 4kb of T-DNA sequence inserted approx. 650 bp after the stop-codon of *PHE1*. In this line we observed significant upregulation of the maternal *PHE1* allele that is not associated with transcriptional activity of promoters localized inside the T-DNA sequence. Upregulation of the maternal *PHE1* allele in SALK_023744 indicates that the region located downstream of *PHE1* might contain an imprinting control region (ICR) and that the repressive influence of this center can be prevented by increasing the physical distance separating the ICR from the promoter of *PHE1*.

We also analyzed whether *PHE1* is upregulated in sporophytic tissues of SALK_023744. Whereas *PHE1* is generally silent in sporophytic tissues, we observed intermediate upregulation of *PHE1* in leaves of SALK_023744 compared to leaves of wild type. This suggests that *PHE1* expression is regulated by two levels – repressing and activating -

and removing the repressing level in the absence of the activation trigger (responsible for active transcription of *PHE1* in developing seeds) causes just stochastic transcription reflected by intermediate mRNA level. Alternatively, we can assume that there are two repressive layers regulating *PHE1* and removing one layer is not sufficient for strong *PHE1* activation.

4.6 Region located downstream of *PHE1* contains DNA-methylation

As previous attempts failed to identify DNA methylation in the promoter and coding regions of *PHE1* (Claudia Kohler, unpublished data), and analysis of *PHE1* imprinting in SALK_023774 indicated a functional importance of the *PHE1* downstream region, we concentrated our investigation on the DNA methylation status at this region. Indeed, we identified a differentially methylated region between wild type and the *met1* methylation mutant. We identified 4 sites of CpG and 1 site of CpTpG being methylated, while all cytosines in an asymmetric context are not methylated. We did not detect significant changes in the methylation status of this region in pollen compared with sporophytic tissues, specifically leaves. Assuming that methylation of the identified sites is indeed functionally important for *PHE1* imprinting, this fact indicates that the default state of this region is to be methylated and there is no pollen-specific activity changing the status of the region during pollen maturation. This leads to the prediction that there should be specific demethylation of this region in the central cell of the female gametophyte, removing DNA-methylation marks on the maternal allele. This prediction would explain why we did not observe an effect on *PHE1* imprinting in *drm1/drm2* mutants. As the endosperm is not genetically transmitted to the next generation, no *de novo* methylation activity would be required to establish *PHE1* imprinting. DEMETER can be considered as one of the candidates for such "DNA-demethylation" activity specific to the central cell (Choi et al., 2002; Kinoshita et al., 2004). A predicted effect of the *demeter* mutation according to this hypothesis should be a significant upregulation of the maternal *PHE1* allele. However, examination of this prediction is complicated by the fact that DME is a positive regulator of *MEA* expression and a direct effect of the *demeter* mutation on *PHE1* expression need to be distinguished from an indirect effect caused by the lack of MEA expression (Choi et al., 2002; Kohler et al., 2003b).

Thus far, we cannot explain how DNA methylation of the region located about 3.5 kb downstream of the *PHE1* promoter region affects its transcriptional status. Many features of genomic imprinting of *PHE1* are analogous to imprinting control of the *Igf2* locus in mice (reviewed in Arney, 2003; Wolffe, 2000). Both loci are maternally inactive and paternally active and both have regions with high DNA-methylation content on the active paternal allele, but also the maternal one it seems for *PHE1*. Furthermore, imprinting of *Igf2* does not depend on non-coding RNAs, similar to our predictions for the *PHE1* locus. Genomic imprinting of the *Igf2* locus in mice is regulated by differential methylation of an insulator element located on the same chromosome several tens of kilobases apart from the *Igf2* gene. The DNA-binding CTCF factor binds only to the unmethylated insulator on the maternally derived chromosome and prevents formation of an activating chromatin loop between the distantly localized enhancer and promoter of *Igf2*. The insulator element on the paternally derived chromosome is methylated, preventing CTCF to bind to the insulator resulting in the formation of an activating chromatin loop between enhancer and promoter of *Igf2* (Murrell et al., 2004). Generalization of the described model is presented in Fig 4-2.

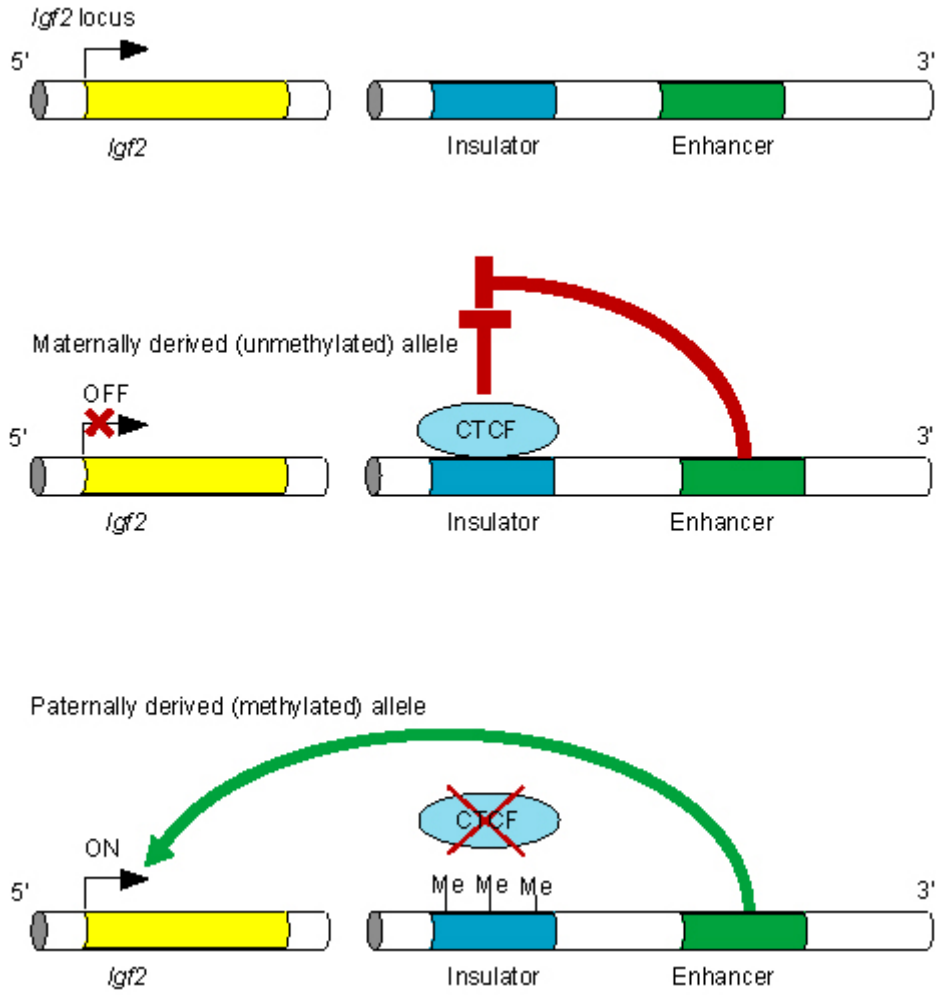


Fig 4-2 | Model summarizing molecular mechanisms of genomic imprinting at the *Igf2* locus (with modifications from Wolffe, 2000). See text for the explanation of the model.

A slightly modified model can be suggested to explain the results of experiments demonstrating importance of the regions located downstream of stop-codon of *PHE1* gene and DNA-methylation for genomic imprinting of *PHE1*. One of the required modifications can be the change of enhancer to silencer and predicting that some important factor FX will bind to methylated insulator and will not bind to unmethylated insulator (Fig 4-3).

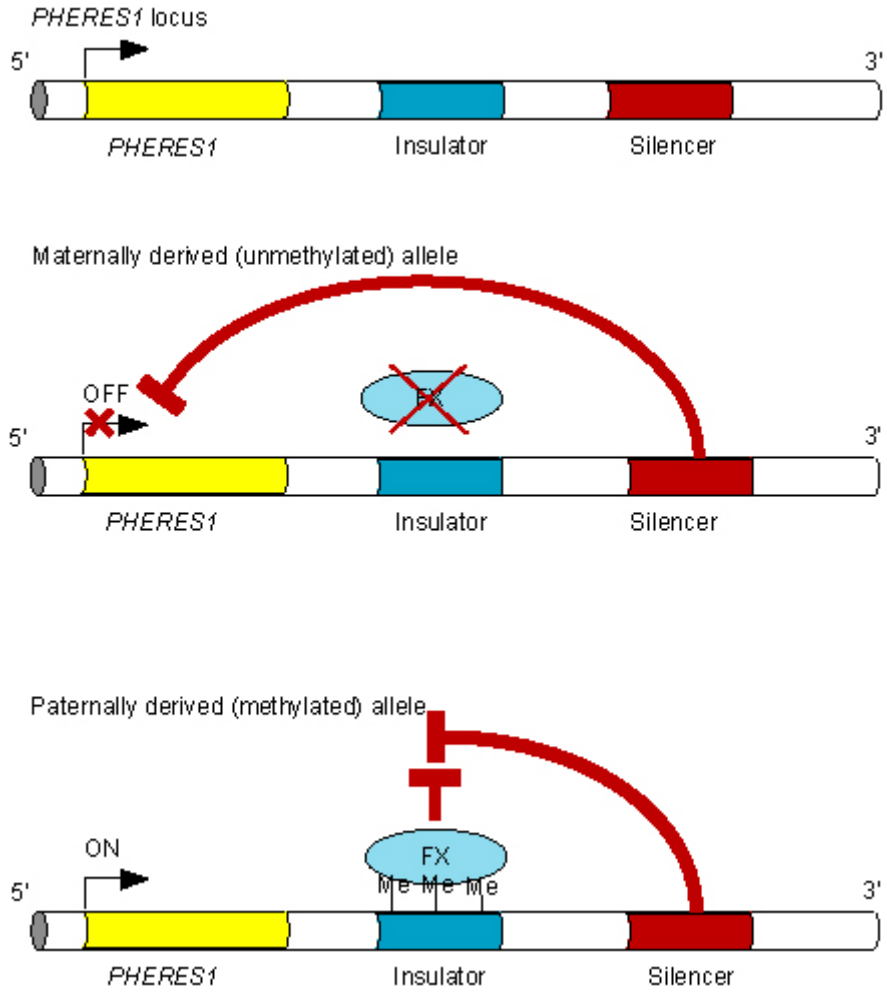


Fig 4-3 | Model explaining molecular mechanisms of *PHE1* imprinting. Similar to the previously presented model explaining genomic imprinting of the *Igf2* gene in mice (Fig. 4-2 and in text), genomic imprinting at the *PHE1* locus in *Arabidopsis* might depend on different chromatin loops formation due to bound/unbound state of the insulator located between *PHE1* promoter and potential silencer. An unknown factor FX binds to the methylated insulator at the paternally derived allele preventing inactivation loop formation between promoter and silencer. The unmethylated maternally derived insulator is not bound by factor FX and in this state is not able to prevent inactivation loop formation between silencer and *PHE1* promoter.

The effect of the T-DNA insertion in SALK_023774 can be explained in the frame of this hypothesis by assuming a breakage of the chromatin loop caused by the 4kb T-DNA or

the formation of a specific insulator-like chromatin structure though the T-DNA preventing interactions between silencer and *PHE1* promoter.

Our hypothesis predicts that the FIS-complex and H3K27 methylation facilitate the formation and stabilization of chromatin loops and therefore, cooperate in imprinting establishment at the *PHE1* locus.

Despite this model explains most of the currently available data concerning *PHE1* imprinting, several evidences are still missing. Thus, analysis of the imprinting status of genetic constructs containing the complete *PHE1* locus, including promoter, coding part, and downstream regions, is required to test the predicted role of downstream regions in imprinting regulation. Furthermore, analysis of the methylation status of *PHE1* in the female gametophyte is required to prove that the identified region is indeed differentially methylated between male and female gametophytes. Analysis of 3D interactions of *PHE1* promoter region with other regions in the genome of *Arabidopsis* using 3C, 4C- or 5C-methods (Dekker et al., 2002; Dostie et al., 2006; Simonis et al., 2006) can provide additional information about localization of the imprinting control region responsible for *PHE1* imprinting.

5 References

Arney, K. L. (2003). H19 and Igf2--enhancing the confusion? *Trends Genet* **19**, 17-23.

Aufsatz, W., Mette, M. F., van der Winden, J., Matzke, M., and Matzke, A. J. (2002). HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. *Embo J* **21**, 6832-41.

Baroux, C., Gagliardini, V., Page, D. R., and Grossniklaus, U. (2006). Dynamic regulatory interactions of Polycomb group genes: MEDEA autoregulation is required for imprinted gene expression in Arabidopsis. *Genes Dev* **20**, 1081-6.

Baroux, C., Spillane, C., and Grossniklaus, U. (2002). Genomic imprinting during seed development. *Adv Genet* **46**, 165-214.

Bartee, L., Malagnac, F., and Bender, J. (2001). Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev* **15**, 1753-8.

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* **427**, 164-7.

Bowler, C., Benvenuto, G., Laflamme, P., Molino, D., Probst, A. V., Tariq, M., and Paszkowski, J. (2004). Chromatin techniques for plant cells. *Plant J* **39**, 776-89.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039-43.

Cao, R., and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* **14**, 155-64.

Cao, X., Aufsatz, W., Zilberman, D., Mette, M. F., Huang, M. S., Matzke, M., and Jacobsen, S. E. (2003). Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr Biol* **13**, 2212-7.

Cao, X., and Jacobsen, S. E. (2002). Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol* **12**, 1138-44.

Cavalli, G. (2006). Chromatin and epigenetics in development: blending cellular memory with cell fate plasticity. *Development* **133**, 2089-94.

Chan, S. W., Henderson, I. R., and Jacobsen, S. E. (2005). Gardening the genome: DNA methylation in Arabidopsis thaliana. *Nat Rev Genet* **6**, 351-60.

Chan, S. W., Zilberman, D., Xie, Z., Johansen, L. K., Carrington, J. C., and Jacobsen, S. E. (2004). RNA silencing genes control de novo DNA methylation. *Science* **303**,

1336.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y. H., Sung, Z. R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**, 5263-76.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J. J., Goldberg, R. B., Jacobsen, S. E., and Fischer, R. L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* **110**, 33-42.

Clark, S. J., Harrison, J., Paul, C. L., and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res* **22**, 2990-7.

Constancia, M., Kelsey, G., and Reik, W. (2004). Resourceful imprinting. *Nature* **432**, 53-7.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185-96.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* **295**, 1306-11.

Desai, U. J., and Pfaffle, P. K. (1995). Single-step purification of a thermostable DNA polymerase expressed in Escherichia coli. *Biotechniques* **19**, 780-2, 784.

Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., Rubio, E. D., Krumm, A., Lamb, J., Nusbaum, C., Green, R. D., and Dekker, J. (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* **16**, 1299-309.

Fuchs, J., Demidov, D., Houben, A., and Schubert, I. (2006). Chromosomal histone modification patterns--from conservation to diversity. *Trends Plant Sci* **11**, 199-208.

Gehring, M., Huh, J. H., Hsieh, T. F., Penterman, J., Choi, Y., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (2006). DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**, 495-506.

Gong, Z., Morales-Ruiz, T., Ariza, R. R., Roldan-Arjona, T., David, L., and Zhu, J. K. (2002). ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* **111**, 803-14.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M., and Coupland,

G. (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. *Nature* **386**, 44-51.

Goto, Y., and Takagi, N. (1998). Tetraploid embryos rescue embryonic lethality caused by an additional maternally inherited X chromosome in the mouse. *Development* **125**, 3353-63.

Grossniklaus, U., Vielle-Calzada, J. P., Hoeppner, M. A., and Gagliano, W. B. (1998). Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. *Science* **280**, 446-50.

Guitton, A. E., and Berger, F. (2005). Control of reproduction by Polycomb Group complexes in animals and plants. *Int J Dev Biol* **49**, 707-16.

Heard, E. (2004). Recent advances in X-chromosome inactivation. *Curr Opin Cell Biol* **16**, 247-55.

Heard, E., and Disteche, C. M. (2006). Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* **20**, 1848-67.

Jackson, J. P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P. B., Cheng, X., Schubert, I., Jenuwein, T., and Jacobsen, S. E. (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. *Chromosoma* **112**, 308-15.

Jackson, J. P., Lindroth, A. M., Cao, X., and Jacobsen, S. E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556-60.

Jeddeloh, J. A., Stokes, T. L., and Richards, E. J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat Genet* **22**, 94-7.

Jullien, P. E., Katz, A., Oliva, M., Ohad, N., and Berger, F. (2006a). Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis. *Curr Biol* **16**, 486-92.

Jullien, P. E., Kinoshita, T., Ohad, N., and Berger, F. (2006b). Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *Plant Cell* **18**, 1360-72.

Kankel, M. W., Ramsey, D. E., Stokes, T. L., Flowers, S. K., Haag, J. R., Jeddeloh, J. A., Riddle, N. C., Verbsky, M. L., and Richards, E. J. (2003). Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics* **163**, 1109-22.

Kantor, B., Shemer, R., and Razin, A. (2006). The Prader-Willi/Angelman imprinted domain and its control center. *Cytogenet Genome Res* **113**, 300-5.

Kermicle, J. L. (1970). Dependence of the R-Mottled Aleurone Phenotype in Maize on Mode of Sexual Transmission. *Genetics* **66**, 69-85.

Killian, J. K., Byrd, J. C., Jirtle, J. V., Munday, B. L., Stoskopf, M. K., MacDonald, R. G., and Jirtle, R. L. (2000). M6P/IGF2R imprinting evolution in mammals. *Mol Cell* **5**, 707-16.

Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S. E., Fischer, R. L., and Kakutani, T. (2004). One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. *Science* **303**, 521-3.

Kinoshita, T., Yadegari, R., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (1999). Imprinting of the MEDEA polycomb gene in the Arabidopsis endosperm. *Plant Cell* **11**, 1945-52.

Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (1999). Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc Natl Acad Sci U S A* **96**, 4186-91.

Kohler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U., and Gruissem, W. (2003a). Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *Embo J* **22**, 4804-14.

Kohler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W., and Grossniklaus, U. (2003b). The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes Dev* **17**, 1540-53.

Kohler, C., Page, D. R., Gagliardini, V., and Grossniklaus, U. (2005). The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. *Nat Genet* **37**, 28-30.

Kroj, T., Savino, G., Valon, C., Giraudat, J., and Parcy, F. (2003). Regulation of storage protein gene expression in Arabidopsis. *Development* **130**, 6065-73.

Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., Higgins, M., Feil, R., and Reik, W. (2004). Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat Genet* **36**, 1291-5.

Lewis, A., and Reik, W. (2006). How imprinting centres work. *Cytogenet Genome Res* **113**, 81-9.

Lindroth, A. M., Cao, X., Jackson, J. P., Zilberman, D., McCallum, C. M., Henikoff, S., and Jacobsen, S. E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* **292**, 2077-80.

Lindroth, A. M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., Jenuwein, T., Khorasanizadeh,

S., and Jacobsen, S. E. (2004). Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. *Embo J* **23**, 4286-96.

Lotan, T., Ohto, M., Yee, K. M., West, M. A., Lo, R., Kwong, R. W., Yamagishi, K., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* **93**, 1195-205.

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E. S., Peacock, W. J., and Chaudhury, A. M. (1999). Genes controlling fertilization-independent seed development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **96**, 296-301.

Mager, J., Montgomery, N. D., de Villena, F. P., and Magnuson, T. (2003). Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat Genet* **33**, 502-7.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Kohler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep* **7**, 947-52.

Malagnac, F., Bartee, L., and Bender, J. (2002). An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. *Embo J* **21**, 6842-52.

Mancini-Dinardo, D., Steele, S. J., Levorse, J. M., Ingram, R. S., and Tilghman, S. M. (2006). Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev* **20**, 1268-82.

Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., and Simon, J. A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**, 197-208.

Murfett, J., Wang, X. J., Hagen, G., and Guilfoyle, T. J. (2001). Identification of Arabidopsis histone deacetylase HDA6 mutants that affect transgene expression. *Plant Cell* **13**, 1047-61.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat Genet* **36**, 889-93.

Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (1999). Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell* **11**, 407-16.

Pandey, R., Muller, A., Napoli, C. A., Selinger, D. A., Pikaard, C. S., Richards, E. J., Bender, J., Mount, D. W., and Jorgensen, R. A. (2002). Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Res* **30**, 5036-55.

Papaioannou, V. E., and West, J. D. (1981). Relationship between the parental origin of the X chromosomes, embryonic cell lineage and X chromosome expression in mice. *Genet Res* **37**, 183-97.

Parenicova, L., de Folter, S., Kieffer, M., Horner, D. S., Favalli, C., Busscher, J., Cook, H. E., Ingram, R. M., Kater, M. M., Davies, B., Angenent, G. C., and Colombo, L. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. *Plant Cell* **15**, 1538-51.

Pauler, F. M., and Barlow, D. P. (2006). Imprinting mechanisms--it only takes two. *Genes Dev* **20**, 1203-6.

Probst, A. V., Fagard, M., Proux, F., Mourrain, P., Boutet, S., Earley, K., Lawrence, R. J., Pikaard, C. S., Murfett, J., Furner, I., Vaucheret, H., and Mittelsten Scheid, O. (2004). Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. *Plant Cell* **16**, 1021-34.

Reik, W., Constancia, M., Fowden, A., Anderson, N., Dean, W., Ferguson-Smith, A., Tycko, B., and Sibley, C. (2003). Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. *J Physiol* **547**, 35-44.

Reik, W., and Lewis, A. (2005). Co-evolution of X-chromosome inactivation and imprinting in mammals. *Nat Rev Genet* **6**, 403-10.

Saze, H., Mittelsten Scheid, O., and Paszkowski, J. (2003). Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat Genet* **34**, 65-9.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., and Goodrich, J. (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *Embo J* **25**, 4638-49.

Schwartz, Y. B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* **8**, 9-22.

Scott, R. J., Spielman, M., Bailey, J., and Dickinson, H. G. (1998). Parent-of-origin effects on seed development in Arabidopsis thaliana. *Development* **125**, 3329-41.

Sharman, G. B. (1971). Late DNA replication in the paternally derived X chromosome of

female kangaroos. *Nature* **230**, 231-2.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet* **38**, 1348-54.

Sims, R. J., 3rd, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends Genet* **19**, 629-39.

Sleutels, F., Zwart, R., and Barlow, D. P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**, 810-3.

Soejima, H., and Wagstaff, J. (2005). Imprinting centers, chromatin structure, and disease. *J Cell Biochem* **95**, 226-33.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* **6**, 846-56.

Sung, S., and Amasino, R. M. (2004). Vernalization and epigenetics: how plants remember winter. *Curr Opin Plant Biol* **7**, 4-10.

Takagi, N., and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256**, 640-2.

Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y., and Feil, R. (2004). Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36**, 1296-300.

Vielle-Calzada, J. P., Baskar, R., and Grossniklaus, U. (2000). Delayed activation of the paternal genome during seed development. *Nature* **404**, 91-4.

Vielle-Calzada, J. P., Thomas, J., Spillane, C., Coluccio, A., Hoeppner, M. A., and Grossniklaus, U. (1999). Maintenance of genomic imprinting at the Arabidopsis medea locus requires zygotic DDM1 activity. *Genes Dev* **13**, 2971-82.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y., and Fuks, F. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871-4.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J. C., Nagy, A., and Magnuson, T. (2001). Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat Genet* **28**, 371-5.

Wilkins, J. F., and Haig, D. (2003). What good is genomic imprinting: the function of parent-specific gene expression. *Nat Rev Genet* **4**, 359-68.

Wolffe, A. P. (2000). Transcriptional control: imprinting insulation. *Curr Biol* **10**, R463-

5.

Wood, A. J., and Oakey, R. J. (2006). Genomic imprinting in mammals: emerging themes and established theories. *PLoS Genet* **2**, e147.

Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., Jacobsen, S. E., and Carrington, J. C. (2004). Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* **2**, E104.

Zemach, A., Li, Y., Ben-Meir, H., Oliva, M., Mosquna, A., Kiss, V., Avivi, Y., Ohad, N., and Grafi, G. (2006). Different domains control the localization and mobility of LIKE HETEROCHROMATIN PROTEIN1 in Arabidopsis nuclei. *Plant Cell* **18**, 133-45.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W., Chen, H., Henderson, I. R., Shinn, P., Pellegrini, M., Jacobsen, S. E., and Ecker, J. R. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* **126**, 1189-201.

Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* **39**, 61-9.

Zwart, R., Sleutels, F., Wutz, A., Schinkel, A. H., and Barlow, D. P. (2001). Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. *Genes Dev* **15**, 2361-6.

6 Acknowledgements

I can not find words to express my deepest feeling of gratitude to my soul-mate Polina for staying with me during all these years.

I am thankful to my parents for the long hours of discussions about what is important in life and what is not, during which I understood that only I can say what is really important for me.

I am thankful to Prof. Claudia Koehler for giving me the opportunity to work in her group towards my degree, for her personal support in critical situations and her understanding of my internal conflicts.

I am thankful to members of my thesis committee Prof. Ortrun Mittelsten Scheid and Prof. Ueli Grossniklaus for their good advices and comments on the thesis

I am thankful to my colleagues in the Koehler group for years spent together.

I am thankful to members of groups of Prof. U.Grossniklaus and Prof. W. Gruissem for discussions and technical assistance.

Curriculum Vitae

Surname: Makarevich

First name: Grigory

Date of birth: December 14, 1979

Place of birth: Novosibirsk, Russian Federation

Citizenship: Russian Federation

Educational history:

1994 – 1996

High school, Novosibirsk, Russian Federation

1996 – 2001 Novosibirsk State University, Department of Biological Sciences, Novosibirsk, Russia M.Sc. diploma thesis title: "Underreplication of DNA in areas of intercalating heterochromatin 11A6-9 and 89E1-2 at *Drosophila melanogaster*".

2004 – 2007 PhD thesis in the laboratory of Plant Developmental Biology supervised by Prof. Dr. C. Köhler, Institute of Plant Science, ETH, Zurich, Switzerland
PhD thesis title: "Regulation of allele specific gene expression of the Polycomb group target gene *PHERES1*"

Work history:

2004-2007 Research assistant at ETH, Zurich, Switzerland

2003-2004 Research assistant at University of Minnesota (UMN), USA

2001-2003 Software developer at "PleskRu", Novosibirsk, Russian Federation

2000-2001 Programmer at "Auktion24", Novosibirsk, Russian Federation

1998-2000 Research assistant at Institute of Cytology and Genetics, Novosibirsk, Russian Federation

Publications/abstracts:

Kohler, C. and **Makarevich G.**(2006). "Epigenetic mechanisms governing seed development in plants." EMBO Rep **7**(12): 1223-7.

Makarevich, G., Leroy O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., Kohler, C. (2006). "Different Polycomb group complexes regulate common target genes in Arabidopsis." EMBO Rep 7(9): 947-52.

Degtyarev, S., Boykova T., Grishanin, A., Belyakin, S., Rubtsov, N., Karamysheva, T., **Makarevich, G.**, Akifyev, A., Zhimulev, I. (2004). "The molecular structure of the DNA fragments eliminated during chromatin diminution in Cyclops kolensis." Genome Res 14(11): 2287-94.

Zhimulev, I. F., Belyaeva E. S., Makunin, I. V., Pirrotta, V., Volkova, E. I., Alekseyenko, A. A., Andreyeva, E. N., **Makarevich, G. F.**, Boldyreva, L. V., Nanayev, R. A., Demakova, O. V. (2003). "Influence of the SuUR gene on intercalary heterochromatin in Drosophila melanogaster polytene chromosomes." Chromosoma **111**(6): 377-98.

Alekseyenko, A. A., Demakova O. V., Belyaeva, E. S., **Makarevich, G. F.**, Kotlikova, I. V., Nothiger, R., Zhimulev, I. F. (2002). "Dosage compensation and intercalary heterochromatin in X chromosomes of Drosophila melanogaster." Chromosoma **111**(2): 106-13.

Moshkin, Y. M., Alekseyenko A. A., Semeshin, V. F., Spierer, A., Spierer, P., **Makarevich, G. F.**, Belyaeva, E. S., Zhimulev, I. F. (2001). "The bithorax complex of Drosophila melanogaster: Underreplication and morphology in polytene chromosomes." Proc Natl Acad Sci U S A **98**(2): 570-4.